

WORKING P A P E R

Using the Logsum as an Evaluation Measure

Literature and Case Study

GERARD DE JONG, MARITS PIETERS,
ANDREW DALY, IRMA GRAAFLAND,
ERIC KROES, AND CARL KOOPMANS

WR-275-AVV

May 2005

Prepared for AVV Transport Research Centre

This product is part of the RAND Europe working paper series. RAND working papers are intended to share researchers' latest findings and to solicit additional peer review. This paper has been peer reviewed but not edited. Unless otherwise indicated, working papers can be quoted and cited without permission of the author, provided the source is clearly referred to as a working paper. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors. RAND® is a registered trademark.

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 00 MAY 2005		2. REPORT TYPE N/A		3. DATES COVERED -	
4. TITLE AND SUBTITLE Using the Logsum as an Evaluation Measure				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) RAND Corporation, 1776 Main Street, PO Box 2138, Santa Monica, CA 90407-2138				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 75	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Preface

The objective of the project ‘Using the logsum as an evaluation measure’ that RAND Europe carried out for the Transport Research Centre (AVV) of the Dutch Ministry of Transport, Public Works and Water Management was to:

Gain experience and weigh the advantages and disadvantages in a case study using both the classic cost-benefit analysis (CBA) approach and the logsum approach for project appraisal.

This report presents the outcomes of both parts of this project, namely:

- A literature review
- A case study using the Dutch National Model System (LMS) for transport.

This report was written for transport modellers and/or economists and other researchers with an interest in the assessment of transport projects.

RAND Europe is an independent not-for-profit policy research organisation that serves the public interest by improving policymaking and informing public debate. This report has been peer-reviewed in accordance with RAND's quality assurance standards (see <http://www.rand.org/about/standards/>) and therefore may be represented as a RAND Europe product.

For more information about RAND Europe or this document, please contact Gerard de Jong.

Prof. dr. Gerard de Jong
RAND Europe
Newtonweg 1
2333 CP LEIDEN
The Netherlands
+31 721 524 51 51
jong@rand.org

Contents

Preface.....	iii
CHAPTER 1 Introduction	1
1.1 Background and objective	1
1.2 Phasing of the project	1
1.3 Contents of this report	2
CHAPTER 2 Introduction to the main concepts.....	3
2.1 Consumer surplus	3
2.2 The logsum measure.....	6
CHAPTER 3 Review of papers on the theory of the logsum as an evaluation measure	11
3.1 The early RUM literature	12
3.2 Income effect and taste variation	13
3.3 Overview papers	14
3.4 Harris and Tanner (1974)	14
3.5 Cochrane (1975)	15
3.6 Daly and Zachary (1976)	16
3.7 Williams (1977)	17
3.8 McFadden (1978)	19
3.9 Ben-Akiva and Lerman (1979)	19
3.10 McFadden (1981)	20
3.11 Small and Rosen (1981)	22
3.12 Börsch-Supan (1990)	23
3.13 McFadden (1996)	23
3.14 Herriges and Kling (1999).....	24
3.15 Karlström (2000).....	26
3.16 Bates (2003).....	27
3.17 von Haefen (2003)	28
3.18 Cherchi, Polak and Hyman (2004)	29
3.19 Daly (2004).....	30
CHAPTER 4 Review of papers on the application of the logsum as an evaluation measure	33
4.1 EXPEDITE consortium (2002).....	33

4.2	Odeck, Rekdal and Hamre (2003)	37
4.3	Castiglione, Freedman and Davidson (2003)	38
4.4	Ecorys Transport and 4cast, May 2004	40
4.5	Gupta, Kalmanje and Kockelman (2004)	41
4.6	Kalmanje and Kockelman (2004)	43
4.7	Koopmans and Kroes (2004)	45
4.8	De Raad (2004)	46
4.9	RAND Europe (2004)	47
4.10	US Department of Transportation, Federal Highway Administration (2004?)	48
CHAPTER 5	Outcomes from the LMS case study	51
5.1	The 'classic' and the 'improved classic' method	51
5.2	The logsum method	56
5.3	Comparing results of 'classic' and 'logsum' methods	63
CHAPTER 6	Summary, conclusions and recommendations	65
6.1	Summary and conclusions	65
6.2	Recommendations	67
References	69

1.1 **Background and objective**

Transport infrastructure projects in The Netherlands are assessed ex ante by using cost-benefit analysis (CBA) procedures following the OEI-guidelines (CPB and NEI, 2000). The project benefits for travellers are incorporated in the form of changes in demand (e.g. from the Dutch national model system LMS or the regional models NRM) and changes in the generalised travel costs (using values of time from Stated Preference studies to monetised travel time savings), and applying the rule of half.

The Transport Research Centre (AVV) of the Dutch Ministry of Transport, Public Works and Water Management has commissioned RAND Europe to perform a study on the use of the logsum (change) as a measure of the change in consumer surplus that results from a transport infrastructure project. RAND Europe has also subcontracted Professor Carl Koopmans of SEO to join the project team.

The objective of this project is to:

Gain experience and weigh the advantages and disadvantages in a case study using both the classic CBA approach and the logsum approach for project appraisal.

1.2 **Phasing of the project**

This project consists of two parts:

1. A review of the literature on the use of logsums as a measure of consumer surplus change in project assessment.
2. A case study with the LMS in which three methods are compared for a specific project (Rondje Randstad: additional high speed and intercity trains connecting the four main cities in the Randstad: Amsterdam, The Hague, Rotterdam and Utrecht):
 - a. the classic CBA approach,

- b. the improved classic CBA approach (following the short term improvements described in Ecorys and 4Cast, 2004),
- c. and the logsum approach.

1.3 **Contents of this report**

This report provides the outcomes of both parts. In chapter 2 the concept of consumer surplus is introduced first, followed by that of the logsum. This chapter gives the existing textbook representation. Subsequent chapters delve deeper into both theoretical issues and applications to provide consumer surplus change in money units. In chapter 3 reviews of papers on the theoretical aspects of using logsums as an evaluation measure are given. Chapter 4 contains reviews of papers on applications of the logsum in evaluation of transport projects and policies, focusing on recent applications to give the state-of-the-art/state-of-practice. The outcomes for the case study using the LMS are in chapter 5. Finally in chapter 6 a summary is provided and conclusions are drawn.

CHAPTER 2 **Introduction to the main concepts**

2.1 **Consumer surplus**

In welfare economics, the level of welfare is measured by adding up the amounts of money that people are willing to pay for goods and services. These amounts reflect the value that goods and services have for people. The ‘willingness-to-pay’ is not equal to what people actually pay. Markets tend to have one price for all consumers (sometimes there are different prices for separate groups; in this case we see one price for one particular group). Given this price, only consumers whose willingness-to-pay is equal to or larger than the market price, will buy the good. The welfare they obtain is equal to their willingness-to-pay. People whose willingness-to-pay is smaller than the market price, will not buy the good. They obtain no welfare from this particular good.

Figure 1 summarises the willingness-to-pay of different consumers in the demand curve. In this curve, all consumers are ranked with respect to their willingness-to-pay, and their individual volumes of demand are added up. On the left side of the curve, we find consumers with a high willingness-to-pay. As we go to the right, we add up the demand of other consumers with a lower willingness-to-pay.

At a given market price, the quantity that is bought is determined by the demand curve. The consumers whose willingness-to-pay is higher than the market price, buy the good. However, their willingness-to-pay is higher than the price they pay. This implies that they derive additional welfare from buying the good. Their welfare is given by the sum of the areas of the shaded rectangle and the shaded triangle. They pay only the area of the rectangle. Therefore, their net welfare is given by the area of the triangle. This net welfare is called the **consumer surplus**.

Figure 1. Demand, willingness-to-pay and consumer surplus

If the price of the goods falls (for instance if transport costs drop as a consequence of government policies), the amount of goods bought generally goes up¹. Figure 2 shows this: the new quantity is higher than the old quantity. The area of the triangle is increased with the heavily shaded areas. Consumers who have already agreed to pay the old price, still have the same willingness-to-pay, but they actually pay less than before. Their welfare change is this cost reduction, the heavily shaded rectangle in Figure 2. Their welfare change is equal to the cost reduction multiplied by 'old demand'.

The price drop induces new consumption: consumers whose willingness-to-pay is smaller than the old market price, but higher than the new market price, decide to buy the good. The net welfare change for these users is their willingness-to-pay minus the price they actually pay (the new market price). This is equal to the area of the heavily shaded triangle. It can be computed as the price change multiplied by the quantity change, times one half (because it is a triangle). This is the so-called 'rule-of-half', which is often used in transport economics.

¹ For so-called inferior goods, the 'income effect' may be negative, and even dominate the positive substitution effect. We do not consider this (improbable) case here.

Figure 2. Changes of market price and consumer surplus

The rule-of-half only applies if the demand curve is (approximately) a straight line. However, the implicit demand curves in transport models are generally not straight lines. Therefore, the rule-of-half may be considered as only a rough approximation of real welfare changes. These real changes can be estimated more precisely by deriving them from the transport models themselves. For some discrete choice models, the change in welfare is defined as the change in the logsum (see below).

The consumer surplus (CS) as defined above was first proposed by Marshall and is derived using the Marshallian or uncompensated demand curve. A price change generally leads to both a substitution effect (from one good to the other) and an income effect (more or less purchasing power for all goods because of an expansion or contraction of the budget). The Marshallian demand curve gives the substitution effect only; the Hicksian or compensated demand curve gives substitution and income effects. Hicks also proposed an alternative measure for the change in welfare, the **compensating variation** (CV).

The CV gives the maximum amount of money (just as the consumer surplus, it is a money measure of utility change) that can be taken from the consumer while leaving him just as well off as before the price reduction (willingness to pay for a price reduction). In case of an increase in the price, the CV is the minimum amount of money that must be given to the consumer to compensate him for the price increase (willingness to accept a price increase).

A related money measure of utility change is the **equivalent variation** (EV). This is the minimum amount of money that must be given to a consumer to make him as well off as he could have been after the price reduction. For a price increase, EV is the maximum amount a consumer is willing to pay to prevent the price increase (Johansson, 1993). So, the CV uses the old utility level and the new prices, whereas the EV uses the new utility level and the old prices.

CV and EV are both areas under the compensated demand curve. These give different outcomes than the CS, which is based on the uncompensated demand curve. In general the CS will be between the CV and the EV. The measures based on the Hicksian demand curve are more attractive, because these do not assume a constant marginal utility of income. In practice the CS is by far the most used measure (sometimes explicitly as an approximation of the CV), because researchers felt that information for CV or EV was lacking (see Cherchi et al, 2004). However, Deaton and Muellbauer (1980) state that there are methods for calculating the CV from price and quantity changes.

2.2 The logsum measure

In this section we provide an introduction to the concept of logsums, largely following the most recent textbook on discrete choice models (Train, 2003). This book focuses on describing the second generation of qualitative choice analysis models with the purpose of bringing these ideas together in a format that makes the methods accessible to a wide audience. The advances have mostly centred on simulation.

A separate section of this book is devoted to describing the calculation of the consumer surplus for policy analysis.

In the field of policy analysis, the researcher is mostly interested in measuring a change in consumer surplus that results from a particular policy.

The consumer surplus associated with a set of alternatives is, under the logit assumptions, easy to calculate. By definition, a person's consumer surplus is the utility, in money terms, that a person receives in the choice situation. The decision-maker n chooses the alternative that provides the greatest utility:

$$\text{Consumer surplus } CS_n = (1/\alpha_n) \max_j (U_{nj} \forall j)$$

$$U_{nj} = \text{the utility that decision maker } n \text{ obtains from alternative } j \text{ } (n = 1, \dots, N ; j = 1, \dots, J)$$

$$\alpha_n = \text{marginal utility of income} = dU_{nj} / dY_n$$

$$\text{where } Y_n = \text{income of person } n, \text{ and } U_n \text{ the overall utility for the person } n.$$

Note that the division by α_n in the consumer surplus formula, translates utility into money units (e.g. dollars, euros) since $1/\alpha_n = dY_n / dU_{nj}$

This utility is known to the decision-maker, but not to the researcher. The researcher observes some attributes of the alternatives as faced by the decision-maker, labelled $x_{nj} \forall j$

and some attributes of the decision-maker, labelled s_n and can specify a function that relates these observed factors to the decision-maker's utility:

$$V_{nj} = V(x_{nj}, s_n) \forall j = \text{"representative utility"};$$

Furthermore, utility is decomposed into an observed and an unobserved (random) component:

$$U_{nj} = V_{nj} + \varepsilon_{nj}$$

where ε_{nj} captures the factors that affect utility, but are not observable by the researcher.

Taking this into account, the researcher is able to calculate the expected consumer surplus by:

$$E(CS_n) = (1/\alpha_n) E [\max (V_{nj} + \varepsilon_{nj} \quad \forall j)]$$

where the expectation is over all possible values of the ε_{nj} 's.

If each ε_{nj} is iid extreme value and utility is linear in income (that is α_n is constant with respect to income), then the expectation becomes:

$$E(CS_n) = (1/\alpha_n) \ln \left(\sum_{j=1}^J e^{V_{nj}} \right) + C$$

where C is an unknown constant that represents the fact that the absolute value of utility cannot be measured.

The term in parentheses in this expression is the denominator of a logit choice probability

$$(P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}}). \text{ Aside from the division and addition of constants, expected consumer}$$

surplus in a logit model is simply the log of the denominator choice probability. This is often called the "logsum term".

Under the standard interpretation of distribution of errors, $E(CS_n)$ is the average consumer surplus in the subpopulation of people who have the same representative utilities as person n. Total consumer surplus in the population can be calculated as the weighted sum of $E(CS_n)$ over a sample of decision-makers, with the weights reflecting the number of people in the population who face the same representative utilities as the sampled person.

The change in consumer surplus is calculated as the difference between the calculation of $E(CS_n)$ under the conditions before the change and the calculation of $E(CS_n)$ after the change (e.g. introduction of policy):

$$\Delta E(CS_n) = (1/\alpha_n) [\ln \left(\sum_{j=1}^{J^1} e^{V_{nj}^1} \right) - \ln \left(\sum_{j=1}^{J^0} e^{V_{nj}^0} \right)]$$

where superscript 0 and 1 refer to before and after the change.

Since the unknown constant C appears in the expected consumer surplus both before and after change, it drops out in calculating the changes in the consumer surplus.

However, to calculate this change in consumer surplus, the researcher must know (or have estimated) the marginal utility in income α_n . Usually a price or cost variable enters the

representative utility and, in case that happens in a linear additive fashion, the negative of its coefficient is α_n by definition.

The formula given in this book for calculating the expected consumer surplus depends critically on the assumption that the marginal utility of income is constant with respect to income. If this is not the case, a far more complex formula is needed, in which α_n becomes a function of the change in attributes.

However, for policy analysis absolute levels are not required, rather only changes in consumer surplus are relevant, and the formula for calculating the expected consumer surplus can be used if the marginal utility of income is constant over the range of implicit changes that are considered by the policy. So, for policy changes that change the consumer surplus by small amounts per person relative to their income, the formula can be used even though in reality the marginal utility of income varies with income.

A slightly different interpretation, namely the logsum as a measure of accessibility, is given in the textbook by Ben-Akiva and Lerman (1985). This book presents the methods of discrete choice analysis and their applications in modelling transportation systems. The authors only briefly describe the mathematical expression to calculate consumer surplus as part of describing a measure of accessibility:

If C_n is a choice set, for multinomial logit:

$$V'_n = \frac{1}{\mu} \ln \sum_{i \in C_n} e^{\mu V_{in}}$$

where V'_n = the systematic component of the maximum utility = measure of accessibility,

μ = scale parameter of the disturbance term ε (ε is by assumption iid Gumbel distributed)

and $E \left[\max_{i \in C_n} U_{in} \right]$ is defined as the measure of accessibility, assuming that the utility scale is established such that $E(\varepsilon) = 0$.

They show that in case of a translationally invariant distribution of the disturbances, the derivative of the accessibility measure (for the multinomial logit model) with respect to the systematic component of the utility of any alternative is equal to that alternative's choice probability. A consequence of the foregoing is that if the distribution of the utilities is translationally invariant then if we increase the utility of every alternative by some amount ΔV , the value of the accessibility measure increases by that same amount. Another consequence is that the derivatives of any two choice probabilities with respect to systematic utilities of the other alternatives will be equal. The effect of a marginal change in the systematic utility of choice alternative i (e.g. due to a price change) on the choice probability of alternative j is equal to the effect of a marginal change of the systematic utility of alternative j on the choice probability of alternative i :

$$\frac{\partial P_n(j)}{\partial V_{in}} = \frac{\partial^2 E \left[\max_{i \in C_n} U_{in} \right]}{\partial V_{in} \partial V_{jn}} = \frac{\partial P_n(i)}{\partial V_{jn}}$$

Thus, we can calculate a measure of consumer surplus. The choice model is viewed as an individual's demand curve for an alternative. The difference in an individual's consumer

surplus between two situations corresponding to attribute vectors x_n^1 and x_n^2 or vectors of systematic utilities V_n^1 and V_n^2 is

$$\sum_{i \in C_n} \int_{V_n^1}^{V_n^2} P(i | V) dV$$

where the choice probability is denoted as conditional on the vector of systematic utilities in order to make the dependency explicit. For the logit model it can be shown that the result of this formula is

$$\frac{1}{\mu} \ln \sum_{i \in C_n^2} e^{\mu V_{in}^2} - \frac{1}{\mu} \ln \sum_{i \in C_n^1} e^{\mu V_{in}^1}$$

which is the difference among expected maximum utilities in the two situations. This measure is expressed in utility terms, but could be transferred to monetary terms in various ways, such as by dividing this measure by a coefficient of travel cost.

More complex discrete choice models (notably Generalised Extreme Value or GEV models) are discussed in the next chapter, especially as part of the summaries of papers by Mc Fadden, who first suggested the GEV family of models).

CHAPTER 3 **Review of papers on the theory of the logsum as an evaluation measure**

This chapter provides an overview of the theoretical literature discussing the issue of calculating overall utility derived by diverse consumers facing a discrete choice and the role of the 'logsum' formula in that calculation. First a general description of the early literature (until the early nineties) is given. Then the current issues (income effects and taste variation) are introduced. After this, the papers are reviewed in alphabetical order.

It is supposed that consumers face a situation in which they must choose one of a finite number of mutually exclusive alternatives². Each alternative has a utility and each consumer chooses the alternative that gives him or her the maximum utility. However, because the consumers are diverse, i.e. have different preferences, the alternative that gives maximum utility may be different for different consumers. Moreover, it is acknowledged that the analyst cannot measure the utilities with perfect precision; any predictions of choice can therefore be made only as probabilities.

This analysis gives the Random Utility Model (RUM) framework, in which consumers (i.e. travellers or freight shippers, in the transport context) are represented as maximising utility, but that this utility is considered random, either because the analyst cannot measure the utilities perfectly or that the consumer does not act consistently, e.g. by making mistakes. This framework has been questioned persistently by psychologists and other social scientists, but remains the only complete paradigm for modelling and evaluating choice behaviour. For the present work we shall remain within the RUM framework.

The theoretical literature reviewed, a total of 15 papers, falls into two phases. First, covering the period up from the early 1970's to the early 1990's, appraisal analysts working within the RUM paradigm constrained themselves to models which did not allow for any effect of income on choice, nor for any variation in tastes which was related to variables in the model. Later, from the mid-90's to the present day, attempts have been made to incorporate these two effects into appraisal models, following successful incorporation of such effect in choice models. It is fair to say that not all of the problems of extending the appraisal models have yet been solved. In any case, practical appraisal procedures up to the present day have almost exclusively been based on the simpler, earlier, models.

² For the purposes of discussion, it is presumed that a market segmentation has been carried out such that consumers within one segment can be considered to choose from the same set of alternatives. This segmentation has the practical consequence that overall utilities must be calculated separately for each segment.

About 20 further theoretical papers have been read for this review but these are not reviewed in detail because they are not directly relevant to the central purpose; informal notes have been prepared on these papers. A small number of further papers could not be reviewed because they could not be obtained within the timescale of the study.

A summary of the theoretical literature review is presented below, split into the two phases.

3.1 The early RUM literature

The key early papers in the RUM literature are McFadden's 1978 and 1981 publications, which form his most important contribution to the discrete choice literature and a major component of his Nobel work. In those papers he first set out the GEV theorem (1978) and then gave full mathematical detail of the links between RUM, choice models and welfare functions (1981) which form the basis for discussing this issue. Essentially, the GEV theorem gives the basis for deriving choice probabilities and overall utilities from a class of functions which satisfy a list of conditions. The specific form of the expression giving the overall utility (the welfare function) is, in simple cases, the log of the sum of the exponentiated utilities of the alternatives, hence acquiring the name 'logsum'.

In many papers, the first publication advocating the use of the logsum as a measure of consumer surplus is stated to be Williams (1977). However, Cochrane (1975) gives the logsum formula for total utility and refers to 1971 work by Neuberger and work parallel to his own by Koenig (neither of these reviewed). Williams himself refers to Neuberger and to Wilson and Kirwan (1969, also not reviewed), in both cases as having used the logsum formula for evaluation. The logsum measure was also in practical use for appraisal before 1977 (by Daly and probably by others, as it is quite simple to derive as the integral of a logit demand function). Both Cochrane and Williams gave a complete theory of utility on which the logsum could be based, but Williams took this further to establish that the logsum was the key 'composite cost' measure which could be used in further modelling to obtain tree (nested) logit models and derived extended logsum measures from tree logit models. McFadden's contribution in this context was to generalise further the models from which logsum-type measures could be derived and to extend and make more rigorous the theory on which their derivation was based.

McFadden's GEV theorem also gives the choice probabilities for the model. These are equal to the derivatives of the logsum with respect to the utilities of the alternatives. That is, the logsum is equal to the integral of any of the choice probabilities with respect to the utility of the corresponding alternative. Given that the choice probability is the expected demand for the alternative from each consumer, it can be seen that the logsum is thus – in some sense – the integral of the ordinary demand curve.

It would thus be convenient to identify the logsum with the Marshallian consumer surplus arising from the choice situation, which is conventionally presented as the integral of the demand curve. However, Marshallian surplus is defined in terms of the integral of demand with respect to the *price* of an alternative, while the logsum is defined as the integral with respect to the *utility* of an alternative. In a context where the marginal value

of money is considered to be constant, this presents no problem. The literature up to the early 1990's, including McFadden, is based on this assumption, which is tantamount to ignoring any influence of income on choice. Other models simply do not deal with the impact of budget constraints on behaviour.

In McFadden's early theory, the key assumptions identifying the models are:

1. the AIRUM assumption (Additive Income RUM), which requires income to enter indirect utility in a specific linear additive form, precluding any income effect on choice behaviour; and
2. the invariant RUM assumption that the distribution of the random component of utility³ is not affected by the values of the observable components – essentially, there is no unobserved taste variation.

A recent paper by Daly (2004) shows that all the key early researchers made these key assumptions (in so far as they discussed the role of income) and also made the same more technical assumptions that are necessary to make the models operational.

3.2 Income effect and taste variation

The impact of income on discrete choice has of course been considered in models of car ownership and other issues for many years, but it seems that McFadden was the first to propose acceptable procedures for calculating consumer surplus measures for models with income effects. There is some confusion about which is the definitive version of McFadden's income-effect work. Hurrigles and Kling refer to a 1995 working paper ('Computing willingness-to-pay in travel demand models'), Karlström to the 1999 publication, possibly revised, with the same title in *Trade, theory and econometrics*; We have reviewed a 1996 version of a different working paper ('On the computation of willingness-to-pay in travel demand models', which refers back to the 1995 paper), circulated at a conference in Stockholm, and the 1998 publication of an altered version of that paper ('Measuring willingness-to-pay for transportation improvements') in the proceedings of the conference, *Theoretical Foundations of Travel Choice Modelling*. The 1996 working paper is the best currently available and that is the text reviewed.

This paper gives three methods for assessing consumer surplus with models that are nonlinear in income: a simulation procedure; an approximation based on a representative consumer approach; and some bounds on the true value of the surplus. Hurrigles and Kling (1999) test these approaches on real data, concluding that the calculation of bounds is inconvenient and may be inaccurate but are unable to choose decisively among the other McFadden approaches and more approximate methods.

However Cherchi *et al.* (2004) conclude that approximations obtained by linearising the demand model may give substantial error. Karlström offers an alternative calculation procedure to replace the McFadden simulation.

³ More strictly, the distribution of the *differences* of the random components.

Taste variation presents a different type of difficulty, in that the valuation attached to attributes of the alternatives are not constant. In particular, the money coefficient may vary randomly, which presents complications of a more fundamental kind. Here we may be concerned with the issue of whether variation is viewed as being between individuals only, or possibly also 'within' a given individual. Von Haefen makes his evaluations without apparent concern for this issue and it is possible that this may be a valid approach. It seems the best conclusion at present is to view the issue as being unresolved.

3.3 Overview papers

Two recent papers offer an overview of the field from different points of view: Bates (2003) and Daly (2004). Bates gives a complete overview of current practice⁴ of transport policy appraisal, relating this to the relevant theory. In particular this paper gives an excellent discussion of the strengths and weaknesses of the 'rule-of-a-half' approximation to consumer surplus calculation. Daly reviews the early theory of RUM modelling, showing that all the important researchers were working on basically equivalent hypotheses, which include a constant marginal utility of money. He then goes on to discuss more recent work which abandons this restriction and to discuss the consequences that the various approaches have for appraisal.

3.4 Harris and Tanner (1974)

This TRRL report is one of the first comprehensive accounts of random utility modelling, aimed at both forecasting and appraisal. They state, however, that the ideas they present 'are by no means original', referring to Quandt (1970)⁵.

The general model is set up with a linear generalised cost, with parameters varying across individuals. Choice probabilities (actually the paper is set up in terms of numbers of trips) are derived by integrating the distribution of the parameters over the sub-space in which each alternative is best, as in standard RUM. Similarly the consumer surplus is derived by integrating the utility ('benefit' in the report) of an alternative over the sub-space in which it is chosen.

They then make the limited specialisation of the general model to one in which individual variation is reduced to an additive difference from a measured cost: 'personal difference' models. In this model the GEV result can be derived

$$\partial CS / \partial x_i = p_i$$

and the 'Hotelling' condition follows

⁴ Perhaps, current British practice.

⁵ Quandt, R. E. *The demand for travel: theory and management*, Heath, Lexington, 1970. This work has not been obtained for review.

$$\partial p_i / \partial x_j = \partial p_j / \partial x_i$$

where x represents the measured cost of the alternative and p is the choice probability (number of trips in the report).

They further show that, given a model defined in terms of demand functions, the distribution of individual variation can be derived from it.

The remainder of the paper shows that models which satisfy a different (logarithmic) form of Hotelling condition can be expressed as ‘personal multiplier’ models and investigates some generalisations and specific cases. The last of these leads them to present a consumer surplus function that is essentially of logsum form and show that the logit choice model follows from this.

This is an important report that did not receive sufficient attention when it was published. However, it appears that its main findings have since been incorporated in more widely read publications.

There is also a more extensive unpublished treatment of the same issues by Harris alone⁶, which gives more rigorous foundations to his theory. This paper has not been reviewed in detail.

3.5 **Cochrane (1975)**

The objective of the paper is to give a basis for the gravity model in economic theory, in contrast to the prevalent theory of the time which was based on entropy.

The initial step is to show that the utility of the best opportunity in a zone can reasonably be taken to follow the Gumbel distribution (this is not proved, but reference is made to a paper by Gumbel himself), since this gives the maximum of a number of distributions with quite general upper-tail properties. The number of opportunities in the zone, provided they are indistinguishable to the analyst, simply adds a constant to the utility equal to the log of the number of opportunities in the zone.

This assumption is then shown to give the logit model of choice probabilities and hence the (singly constrained) gravity model. Then “As an interesting aside, we can calculate the total surplus arising from trips actually made.” This is the logsum formula, summed over origins (equation 23).

The doubly constrained gravity model is presented as being an amendment to the singly constrained model with a ‘shadow price’ in each over-subscribed zone which reduces utility. This effect is shown also to influence the utility by entering the logsum.

The unconstrained model is presented as embodying a choice not to travel if the total surplus is inadequate. However, because the utility of not travelling is taken to be constant (zero) rather than also being distributed, the formula derived (equation 34) is more

⁶ Harris, A. J., *Demand for modes of transport: a mathematical theory of personal choice among economic alternatives*, Transport and Road Research Laboratory, Draft Supplementary Report, 1974/5?.

complicated than a modern version would be and it is not obvious how a surplus over the choices of travelling and not travelling should be calculated.

Further sections discuss the relationship of these models to the Intervening Opportunities model, introduce a particular form of taste variation and begin to discuss the issues arising when modes of travel are considered.

3.6 **Daly and Zachary (1976)**⁷

The stated objective of the paper is to search for increased flexibility in modelling choice.

Section 1 sets up the utility maximising framework on which choice is assumed to be based and introduces randomness which is postulated to arise only because of limitations on observation. Consideration is restricted to invariant RUM models (described as ‘location parameter’ models), which are stated to be the only class of models consistent with then current UK generalised cost appraisal procedures, although they claim any RUM model has a benefit measure. They note that Daly and Zachary (1975)⁸, which uses random parameters models, is RUM but not IRUM.

Section 2 presents Zachary’s theorem which is the main contribution of this paper to the literature. This states that, given a probability model defined over alternative utilities, the conditions:

1. the probabilities are unaffected by the addition of a constant to the utilities of all the alternatives;
2. every alternative is potentially dominated, i.e. its probability tends to zero as the utility of any other alternative tends to plus infinity;
3. the probabilities are non-negative and sum to 1;
4. the mixed partial derivative of the probability of an alternative with respect to the utilities of other the other $n-1$ alternatives are non-positive for odd and non-negative for even n ;
5. (Hotelling) the Jacobean of the probabilities with respect to the utilities is symmetric;

are necessary and sufficient for the model to be consistent with invariant RUM. Further, the cumulative distribution of the random terms in the RUM is defined by the probabilities, which in turn implies that any one of the probabilities defines all of the others, as had been proved by Harris and Tanner (see below).

Two important examples are given: binary probit and binary logit.

⁷ This paper was published (with printing errors) in D.A. Hensher and M.Q. Dalvi (eds.), *Determinants of Travel Choice*, Saxon House 1978 (US editions appear to be 1979).

⁸ Daly, A. J. and Zachary, S. (1975) *Commuters’ values of time*, LGORU Report T55, Reading.

Section 3 discusses the MNL model in this context, drawing attention to the distinction between the *numéraire* scale on which utility is defined (and to which condition 1 applies) and scales that may be estimated without imposing constraints across the alternatives. The limitations of the MNL are discussed and the composite utility measure – the logsum – is introduced.

Section 4 shows that the nested logit model, defined by exploiting the logsum measure, is consistent with invariant RUM by applying Zachary's theorem, the conditions for which are satisfied providing the model scale is not greater at the higher level than at the lower level. It is interesting to note that if the condition fails, then the model does not exist, i.e. there is no distribution of the random terms that yields a model whose scale parameters are in the wrong relationship to each other. The generalisation to multiple nests and multiple levels is noted

Section 5 discusses estimation issues, concluding that maximum likelihood is a reasonable method, yielding estimates and confidence limits and that this can be achieved by a general non-linear optimisation routine. Sequential estimation is not ideal, but is consistent.

Section 6 presents an application, showing that the nesting gives a significant improvement in the model.

Section 7 gives simple conclusions from the work.

3.7 Williams (1977)

This landmark paper discusses consistent demand models and evaluation measures in the context of travel demand. It does so in the context of multiple dimensions of demand and on the basis of random utility theory.

Section 1 introduces the area of discussion and lays emphasis on the “sadly neglected” area of economic evaluation. It then presents the main conclusions:

1. that the models currently in use in transport studies were, without exception, erroneous;
2. that recent research on probability theory as a basis for travel demand modelling was largely incorrect;
3. that the use of the ‘rule-of-a-half’ (roah) was computationally inefficient and inappropriate for new facilities.

Williams was clearly not satisfied with the then current approaches! However, the main value of the paper in the longer term has been the constructive alternative methods he proposed.

Section 2 reviews a number of existing models, focussing on the ‘composite cost’ measures used to integrate the mode split and distribution models, then on the user benefit. The Hotelling condition is indicated as the guarantee that a unique calculation of user benefit can be obtained for changes in generalised transport costs. The roah calculation can be

considered to be a rough approximation to this. A range of issues concerned with these procedures is set out for discussion in the remainder of the paper.

Section 3 discusses composite cost measures, showing that current formulae fail to satisfy essential requirements.

In Section 4, Williams contrasts the ‘constant utility’ approach to choice modelling of Luce and others, with its attendant problem of dealing with close substitutes, with the random utility approach. The approach is chosen for RUM in which ‘representative’ (measured) and random (unmeasured) parts are additively separated. The work of Harris and Tanner and of Cochrane is then quoted to derive the basic properties of MNL. Examples are given but in this section the models remain single-level, whether logit or probit. In discussing the quite close relationship of these models with those derived from entropy considerations, Williams shows that the change in surplus is equal to the change in average cost plus the change in entropy divided by the model scale parameter⁹. However, for the remainder of the paper he uses the utility approach on the grounds that a rational decision framework is more appropriate.

Section 5 presents models of multi-dimensional choice, beginning with a description of the choice dimensions of interest and how these have been represented in some of the literature. A key step is to assume separability of the utility over the dimensions of choice (this would not now be seen as essential). Referring to Ben-Akiva¹⁰, sequential models of mode and destination choice with alternative sequences and with a range of ‘composite cost’ formulations are discussed. Apparently, Ben-Akiva discussed statistical tests to show whether one structure was preferable to another, but was not aware (in his thesis work) of the utility interpretations of these structures.

The section goes on to discuss the red bus / blue bus conundrum in the context of correlated attributes and sets up utility functions which make these attributes – both observed and unobserved – explicit. Using utility functions with this separability, Williams develops choice models *and surplus measures* which apply for models with general tree structures. Using logit-type formulations, he produces choice probabilities *and extended logsum measures* (equations 167 and 181) for multi-level choices. The condition on the scale parameters at different levels is derived as a consequence of the addition of independent random variables. The section concludes with speculation about appropriate methods for determining which structure might be appropriate in specific circumstances, including the possibility of cross-nested structures, but without reaching clear conclusions.

Section 6 discusses a number of practical examples, retreating to some extent from the initial statement that all models in current use were incorrect. Finally, Williams concludes, making the important point, *inter alia*, that the form of composite cost measure that is appropriate for a given utility structure is fixed by that structure.

⁹ Entropy is defined as $W = -\sum_j p_j \log p_j$, where p are the choice probabilities. Note $W > 0$.

¹⁰ Ben-Akiva, M. (1973) *Structure of passenger travel demand models*, PhD Dissertation, MIT; not reviewed.

3.8 **McFadden (1978)**

This paper is introduced by emphasising the fact that consumers differ in their tastes and that practical models of the choice of housing location need to recognise this. The elementary alternatives, dwelling units, are grouped into zones and share some observed and unobserved attributes as a consequence of that grouping.

The model is set up as RUM, with a ‘strict utility’ which is measurable but contains unknown parameters and an unknown term summing unobserved variables. Among such models is the MNL, derived when the unknown term is independent and identically distributed Gumbel, but this assumption is unreasonable. A generalisation, the nested logit model, is obtained by representing the total utility of the dwellings within a community as a logsum and allowing that logsum to have a coefficient not equal to 1 in the choice of community.

Section 5 of the paper states and proves the GEV theorem (described in detail in the review of McFadden, 1981). Here we can note that the requirement concerning zero arguments is not stated in the 1978 work. More importantly, the issue of AIRUM is not raised – the assumption is only that utilities are distributed GEV. Following the proof of the theorem and its corollaries, the examples of MNL and nested logit are given, with the implication that the nests need not be disjoint. It is shown that nested logit can be used as a specialisation of GEV for modelling residential choice. The restriction on the nesting coefficients is also derived.

The following section (7) deals with the issue of sampling of alternatives, which is of considerable practical importance for practical discrete choice modelling but not relevant for the present work. It presents the Positive and Uniform Conditioning properties and shows that drawing samples of alternatives that satisfy either of these properties can be used (with appropriate correction in the Positive Conditioning case) to estimate unbiased models. The final main section discusses the issue of the aggregation of alternatives with equal observed variables, leading to what would now be called size or attraction variables. However, McFadden relies primarily on previous literature for this discussion.

3.9 **Ben-Akiva and Lerman (1979)**

This paper discusses the measurement of accessibility as the maximum utility available from a travel choice and presents formulae, including generalised logsums, to measure it.

The paper starts by presenting the widely-used Hansen measure of accessibility and shows that it appears to be based on a naïve understanding of travel choice¹¹. It then suggests that accessibility should be measured as the maximum utility obtainable from a travel choice for a given individual, but without giving any detailed explanation for that choice. In a random utility model, where utility is not known precisely, a reasonable alternative is

¹¹ However, the measure finally derived is in fact equivalent to a Hansen measure!

to use the expected maximum utility. “It is worth noting that under a set of restrictions on the distribution of [utility] the expected value of the maximum utility has a direct interpretation as a consumer surplus measure.” That is, accessibility and consumer surplus are the same thing: the maximum expected utility.

The authors then investigate the conditions under which two convenient properties hold:

1. *that the value of maximum utility increases with the size of the choice set;*

This property holds in all cases, providing the maximum utility can be calculated. It follows that the maximum utility is greater than the utility of any one of the alternatives.

2. *that the value of maximum utility increases with the mean utility of each of the alternatives.*

This property does not hold in general but requires conditions; sufficient conditions are attributed to Harris and Tanner (1974) and to Williams (1977). This requirement is essentially that the utility of an alternative must always increase when the mean value increases, a condition not necessarily satisfied by models with taste variation but always satisfied by (translationally) invariant RUMs.

For the logit model, it is then shown that the accessibility measure is simply the logsum, with the possible addition of Euler’s constant.¹² For the probit model, however, a ‘closed’ form for an exact calculation does not exist but an approximate solution is suggested.

The paper goes on to discuss hierarchies of choice, considering 10 travel decisions, ranging from location of home and work to route choice. It states that considering all choices together would be impossible and therefore breaks choices into mobility and travel blocks, which can be modelled together, although sequential estimation would still be necessary. However, the accessibility (maximum utility) that might apply in hierarchical choice models is not stated clearly although it is implicit as the denominator of the choice formulae given.

3.10 **McFadden (1981)**

McFadden introduces his work (sections 5.1 to 5.3) by discussing how randomness can enter models of behaviour, either through subject variation (e.g. mistakes) or through observer limitations. The key motivation is that differences between subjects must be treated explicitly, rather than modelling in the aggregate.

Then he introduces the notion of a probability choice system (PCS) which depends only on observed attributes of alternatives and the subject. Section 5.5 extends this to random utility models (RUM), which induce a PCS providing the utility distributions can be

¹² Euler’s constant, usually denoted by γ , is approximately 0.577 and is the mean value of a standard Gumbel distribution.

integrated etc.. It is not possible to identify separately taste variation within and between subjects.

The first key section (5.7) sets the discrete choice problem in the context of a classical micro-economic framework and shows that Roy's Identity can be used to obtain choice probabilities. Then a key step is to assume (additive income RUM, AIRUM) that indirect utility can be formulated as (equation 5.12, simplified):

$$V = ((y - q) - \alpha) / \beta$$

with y and q as income and alternative price, α and β homogenous of degree 1 in all prices. Then income falls out of the choice process, although we are warned to be careful of its possible index role.

The second key section (5.8) proves that, assuming AIRUM and under reasonable further conditions, any one of a social surplus (SS) function, a RUM or a TPCS (translationally invariant PCS) implies the other two. This equivalence is described as the Daly-Zachary-Williams theorem and corollaries to it. The assumption of AIRUM should be noted here since it does not appear to be necessary *to the theorem* that income or price should appear at all or in any specific way. Indeed McFadden discusses alternative *numéraires*.

Then, alternative model forms are discussed: Thurstonian (probit), Lucean (logit) and Tverskian (elimination by aspects (EBA)).

The third key section (5.15) presents the GEV theorem. Given a function H satisfying the GEV conditions:

1. H is a non-negative linearly homogeneous function of non-negative arguments;
2. H goes to infinity with any of its arguments;
3. partial derivatives of H exist and are continuous, with non-positive even and non-positive odd mixed derivatives;
4. adding alternatives corresponding to zero arguments of H does not change H .

Then H induces a 'GEV' distribution, say F . For F it follows that

1. a RUM with (utilities – prices) distributed according to F satisfies AIRUM and has an SS equal to $\log H$;
2. the choice probabilities are equal to the differentials of H .

The theorem is proved in McFadden (1978).

In the simplest case, when H is the sum of its arguments, the RUM generated is MNL and the SS is the simple logsum. More general H functions can induce nested logit etc. and then the logsum-like formulae become more complex. This theorem can be taken as the theoretical basis for the logsum measure, but note the AIRUM assumption, which is unnecessary here as it was in the D-Z-W theorem.¹³

¹³ The AIRUM assumption is omitted in the 1978 paper. Further, it is useful to note that, even with the generalisation of GEV to include functions homogeneous of any degree $\mu > 0$ by Ben-Akiva and Francois

The remainder of the main part of the paper: discusses the similarity of MNP, GEV and EBA models, concluding that they are very similar;¹⁴ discusses estimation procedures, concluding that sequential estimation is consistent but may be inefficient; and presents an example from McFadden's work in the Bay Area. Appendices discuss scaling normalisations and give formulae for the various models; present a proof of the D-Z-W theorem; and present the Elimination-by-Strategy model.

3.11 Small and Rosen (1981)

This work aimed to show that “conventional methods of applied welfare economics can be generalized to handle cases in which discrete choices are involved” (p. 106). They note that the switch from continuous to (partially) discrete choice involves a *generalisation* of the conventional methods.

Section 2 gives an up-to-date review of the classical theory of consumer surplus, pointing towards compensating variation as the key measure. The integral of the compensated demand function gives the change in the expenditure function (equation 2.12).

At this point we note that for compensating variation we need to integrate the compensated demand function, whereas we normally have the ordinary demand function. A correction can be made using the Slutsky equation, but this is only approximate unless all consumers have the same income elasticity and each individual's share of aggregate income remains constant under all price and income changes.

Section 2 goes on to discuss the implications in the classical context of taxation, which is one of the main objectives of the paper. The concept of “excess burden”, the welfare loss brought about by distortions in relative prices is introduced, and the ways in which this depends on the use made of tax revenues are discussed.

Section 3 introduces discrete choice, which is described by requiring the consumer to choose none or one of the two goods on offer, i.e. $x_1 \cdot x_2 = 0$. Note that the amount of the good that is chosen is not fixed, i.e. this is a discrete-continuous choice.

Theorem 1 states that the imposition of the discrete choice constraint does not affect the calculation of compensating variation. Equation 3.5 is a simple extension of 2.12. The crucial step in the proof is to note that the compensated demand function can be integrated across the point at which the consumer switches from x_1 to x_2 , because at that point both left and right derivatives exist.

(1983, unpublished and not reviewed), if the lower level utilities have a constant scale relative to money, as in AIRUM, then the logsum will have the scale of μ times that constant relative to money.

¹⁴ But see Batley, R. and Daly, A. (2003) Establishing equivalence between nested logit and hierarchical elimination-by-aspects models, presented to European Transport Conference, Strasbourg.

An important corollary is that by summing over all the individuals we obtain the result that the aggregate compensating variation is the integral of the aggregate compensated demand function.

A further important corollary is that the result can also be applied to expectations, when we use the expected demand, i.e. the product of the compensated choice probability and the conditional compensated demand. However, we need to note here that the function to be integrated is compensated demand, not ordinary demand – this implies that we cannot use results such as the logsum (which is the integral of the ordinary demand function) to obtain compensating variation when this is different from Marshallian surplus. An approximation based on the Slutsky equation can be used.

In passing (footnote 17), the authors claim that the approaches of random utility and random choice based on fixed utility are operationally equivalent.

There follows a discussion of the impact of taxation.

Section 4 discusses the calculation of consumer surplus when the quality of an alternative changes and concludes (Theorem 2) that the same methods can be used.

Section 5 discusses the application of these methods to specific demand models, assuming that the random term is independent of prices, qualities etc., i.e. that we have an invariant RUM in the sense of the rest of the classical papers. Then, assuming there is no income effect, we can derive equation 5.5 which looks much like the earlier integrals (presumably, if there is income effect, the Slutsky correction does not work?).

For the probit model this gives a formula that is simple to write if not to evaluate (equation 5.7) and for the logit model it gives the logsum (equation 5.9). Section 6 gives very brief conclusions.

3.12 **Börsch-Supan (1990)**

This paper examines the necessity of the assumption that structural parameters in a tree logit model should indicate an increase of variance of utility going ‘up’ the tree. It concludes that the assumption is too restrictive. Although the paper is a key part of the literature on choice modelling, it does not discuss the issue of consumer surplus and is not directly relevant to the present review.

3.13 **McFadden (1996)**

The paper aims at presenting methods for calculating willingness to pay for improved alternatives that are chosen in a discrete choice context. The analysis is undertaken in the context of RUM modelling and the example of fishing sites is used.

The indirect utility function for an alternative site would typically include terms for ‘disposable income’ minus costs, travel and recreational time, together with hedonic

attributes and a random term. Allowance is made in the model for the different perception of different types of cost. The restrictive aspects of the typical RUM are additive separability and linearity in both income and hedonic attributes. Invariance of the random term from income or hedonic attributes is a possible further restriction. These restrictions may impact more on the welfare calculations than on demand modelling.

With these restrictions, income ‘drops out’ of the comparison of the alternatives and there are also no income effects in consumer surplus, so Hicksian and Marshallian measures coincide. A social preference function can be derived, which can generate demand functions via Roy’s Identity. But “even additive non-linear income effects, or additive linear income effects with coefficients that differ by alternative, are sufficient to upset ‘log sum’ formulas for WTP”.

To deal with this situation, McFadden develops bounds, based on the income change required to keep the utility of each alternative constant. He shows that (in the case when no alternative deteriorates) the expected utility change is bounded by the averages of these utility changes per alternative, weighted by the original (lower) and final choice probabilities (upper bound). These bounds do not depend on income entering the model linearly.

When income enters linearly, average willingness to pay can be represented as the willingness to pay of the average consumer. However, when income does not enter linearly, an approximation assuming linearity can lead to bias – McFadden suggests circumstances where the bias can reach 30%.

For GEV models and for models composed of mixtures of GEV models (such as mixed logit), McFadden notes that these can approximate any RUM to any desired degree of accuracy. Based on three theorems, he shows that, providing utility is linear and additive in income, a ‘mixed logsum’ formula can be used and that this coincides with traditional Hicksian and Marshallian measures of consumer surplus. With nonlinear income in the utility function, McFadden sets out a simulation procedure, whose computational tractability is subject to discussion. The two key points are the facility with which draws can be made from the mixing density and the solution to find the equivalent or compensating income for each set of draws. McFadden gives some results, showing that it can be difficult to achieve accurate simulation under some circumstances.

3.14 **Herriges and Kling (1999)**

The paper aims at relaxing the constraint of linear income in the indirect utility function, which many previous authors had assumed to avoid difficulty. The assumption of constant marginal utility of income has often been made even in cases when nonlinear functions could reasonably be expected. The assumption has been made because of the difficulty of estimating non-linear demand functions but also because of the difficulty of

computing welfare measures with nonlinear income. Further, linear income allows the approach of a ‘representative consumer’ to be used; they refer to McFadden (1995)¹⁵.

The development of welfare measures in discrete choice is largely ascribed to McFadden, with most applications in transportation although more recently other fields have been studied. This theory assigns to each of the discrete alternatives chosen under an overall income constraint an indirect utility depending on income remaining (my nomenclature), the characteristics of the alternative and a term representing unobserved variables. Compensating variation can be calculated relatively easily in a system of this type if it is assumed that:

- A1 the unobserved term is additive;
- A2 the disturbances are GEV;
- A3 income has a constant marginal utility.

The objective of the paper is to study the relaxation of assumption A3. There are three possible approaches.

The first approach is by simulation, following the procedure suggested by McFadden (1995/6). The problems with this approach are that it requires repeated sampling from the distribution assumed for the unobserved term and that the equation for compensating variation has to be solved for each draw, making the procedure computationally intensive (but see Daly, 2004).

The second approach is to approximate the calculation by using a representative consumer approach. However, this may be inaccurate because of the nonlinearity of income in the model: essentially it calculates the income required to equate average utility before and after the change, not the average of the income required to equate disaggregate utilities. McFadden (1995/6) finds that this approach can be quite inaccurate for large changes.

The third approach is to rely on theoretical bounds, also due to McFadden (1995/6). These can be computed – in both theoretical and ‘computable’ forms – but a number of further difficulties are raised. Further, bounds, which are themselves random numbers dependent on parameters estimated from sample data, are not ideal in practical analysis.

Herriges and Kling analyse data on sport fishing in California, specifying and estimating a number of demand models and using these to calculate welfare benefits of hypothetical changes in the characteristics of fishing alternatives. These are then used to calculate benefits using each of the proposed methods. They conclude that the bounds approach is difficult to implement, since the upper bounds in particular are difficult to estimate. However, the representative consumer and an approach based on approximating the demand model by a linear-in-income version appear on this data to give similar estimates to those obtained from the more theoretically correct sampling approach (but see Cherchi, Polak and Hyman, 2004).

¹⁵ See the discussion in section 1.2. All the points referenced by Herriges and Kling are mentioned in the 1996 McFadden paper which is reviewed.

3.15 Karlström (2000)

The paper presents and applies generalised logsum formulae applicable to the case where the marginal utility of money is not constant. He argues that such cases may be important in practice, not only in developing countries. ‘Representative consumer’ approaches do not work when the marginal utility of money varies across alternatives, *even if there is no variation across consumers*.

Compensating variation for a single consumer is defined as the income change that will return the consumer to the same utility level that he or she experienced before some change in prices or qualities of alternatives. In the random utility context, it is assumed that the random component of utility is unchanged by the changes in the alternatives. The randomness however introduces complications in the calculation of the surplus.

In GEV models, when the marginal utility of money is constant, the logsum measure can be used and the Hicksian and Marshallian measures coincide. However, when the marginal utility of money is not constant, this approach cannot be used. McFadden’s simulation approach is seen as computationally intensive, so that an improved formula would be desirable and that is what is presented in this paper.

The following summary of Karlström’s procedure is taken from Daly (2004).

In a simple case, suppose the only change is in the price x_i^0 of i , which is increased to x_i^1 . If i is not chosen either before or after the change there is no impact. If i is chosen both before and after then the solution for the change in surplus is simply $S = (x_i^1 - x_i^0)$, applying to that part of the population that chooses i after the change. The remaining contribution is for those people who switch away from i (no-one switches to i) who have to be compensated by the amount of price increase that would cause them to be indifferent between i and the best alternative.

Karlström’s exact method to calculate the surplus is based on the function

$$\Pi_i(y) = \Pr \{ (\text{choose } i \text{ before}) \cap (\text{compensated income after} \geq y) \}$$

$\Pi_i(v_{ii})$ gives the compensated choice probability, where v_{ii} is the expenditure level needed to restore income to its initial level (i.e. $y + S$), given that the choice remains i (i.e. $v_{ii} = y + x_i^1 - x_i^0$). For the people who remain choosers of i , the compensated income after the change is v_{ii} , so that $\Pi_i(v_{ii})$ is the probability of choosing i before and after. The surplus for these people is then $(v_{ii} - y) \Pi_i(v_{ii})$.

For the people who switch choices when the compensated income needed is between z and $z + \delta z$, we find that the compensated income they need is z and the number of them is given by minus the derivative of Π (Π declines with y), i.e. the surplus is minus $(z - y) \Pi_i'(z) \delta z$.

Hopefully this gives some insight into Karlström’s main result, which is that the expected mean compensated income for a price change for all the alternatives is given by

$$E v = \sum_i \{ v_{ii} \Pi_i(v_{ii}) - \int_{\underline{\mu}}^{v_{ii}} z \Pi_i'(z) dz \}$$

where $\underline{\mu}$ is the minimum μ_{ik} over the alternatives k . This is claimed to be easy to calculate when the choice model is GEV and to form a generalised logsum.

The formula, once understood, can be applied in any additive RUM model, e.g. in a probit model. However, when income effects are present, Hicksian surplus cannot be presented as the integral of conventional demand functions. Karlström then devotes considerable attention to showing that Slutsky symmetry does not necessarily hold to establish the impossibility of integrating a demand function to derive the surplus.

Karlström then presents two examples to illustrate the theory. One uses the same data as Herriges and Kling (1999), the other uses commuter mode choice data from the Swedish national travel survey. He recommends that further research should focus on extending model forms that can exploit nonlinear income terms.

3.16 **Bates (2003)**

This is a lengthy document, commissioned as a resource paper for the IATBR Conference, which aims to discuss economic evaluation and how the theory and practice can be reconciled. It does not discuss logsums in any detail but gives important background information about alternative procedures (i.e. rule of a half) and the application context.

The first chapter gives a history of evaluation and explains the concept of consumer surplus.

Chapter 2 presents an exposition of the relevant aspects of microeconomics, set out quite briefly. Here the key issues are identified as ‘Slutsky symmetry’ and the separability of income and price. Compensating and equivalent variation are defined and contrasted with Marshallian consumer surplus. The conditions under which a representative consumer approach can be taken, based on Gorman (not reviewed here) which requires ‘quasi-linear’ separability.

The impact of discrete choice on evaluation is to focus attention on *conditional* indirect utility, which is extended by an unknown random term to introduce heterogeneity in the population. McFadden’s 1981 analysis introduces the AIRUM assumption that income enters indirect utility linearly with a uniform coefficient, which allows simple calculation, but Bates notes that more recent work has gone beyond this assumption to more general models.

Chapter 3 discusses the practice of evaluation using the rule of a half. Bates shows that this approximation can be very good in some circumstances. In the UK, a link-level procedure (COBA) was based on this, but a link level procedure can only be used if the OD matrix is fixed. For variable OD-matrices, a new procedure (TUBA) had to be written. Further the rule of a half cannot be used at all when:

- the set of alternatives changes between ‘before’ and ‘after’, e.g. when new modes are introduced (because the ‘before’ cost is effectively infinite);
- there is switching between alternatives which have different average levels of utility and the scale of the switching model becomes large or ultimately when the models becomes deterministic, e.g. in the case of many models of route choice (because the linear approximation no longer holds).

The implication would be that other evaluation methods – e.g. logsums – should be used in these cases, but Bates does not draw this conclusion explicitly.

Further issues arise in partitioning the benefit, e.g. into time and cost savings, or to attribute it to different groups of travellers.

Quite often it is required that the components of utility are ‘re-weighted’, i.e. assigned different coefficients in a rule-of-a-half calculation from their weighting in forecasting demand. In particular, different values of time are often used. In principle this approach is understandable but can lead to inconsistencies. These inconsistencies can be avoided by the use of more sophisticated approaches, but the data and analysis requirements for these approaches are likely to be burdensome in practical situations. Bates also discusses further practical issues.

Bates’s final main chapter deals with research issues: dealing with new modes, treating land-use changes, ‘shadow prices’ such as balancing constants, income effects, where he reports the literature very briefly and finally he revisits the issue of aggregation.

3.17 von Haefen (2003)

Von Haefen “develops an approach to welfare measurement from random utility models (RUMs) that conditions on an individual’s observed choice”. The approach is related to work by Revelt and Train which estimates ‘part-worths’, i.e. components of the utility function, conditional on the observed sequences of choices of individual consumers.

For the present study, conditioning on observed choice may not be of great importance, because often forecasts will be for large populations and alternative scenarios over an extended period, so that current choices of a sample are not very relevant. The relevance of the paper is more in the treatment of taste variation, but conditioning on observed choices may be useful in some transportation contexts.

Von Haefen explains the theory of conditional valuation relative to unconditional valuation (citing Small and Rosen, 1981). He then discusses his data, which relates to water-based recreation on the lower Susquehanna River in Pennsylvania. Models are estimated which allow different levels of variation between individuals and within each individual’s sequence of observed choices. These models are then used to estimate welfare changes for two policy scenarios.

The interest of the paper for the current work is in section 6 which gives procedures for calculating Hicksian welfare measures. In equation (10) he gives an unconditional welfare measure which allows for taste variation *including random variation in the marginal utility of money*. Train (1998, not available for review¹⁶) has suggested a sampling procedure to evaluate this measure. Analogous measures are presented for conditional welfare measures,

¹⁶ Train, K. (1998) Recreation demand models with taste differences over people, *Land Economics*, 74, pp. 230-239.

but these present greater challenges for evaluation, for which computationally intensive Metropolis-Hastings procedures are given by von Haefen.

The paper concludes with calculation of the benefits given by conditional and unconditional measures for each of the four model specifications for each of the two policy scenarios. Von Haefen concludes that the differences are sufficient to suggest that conditional measures may be significantly better than unconditional ones in this type of application.

3.18 **Cherchi, Polak and Hyman (2004)**

The paper is intended to look at the impact of various sophisticated aspects of travel demand models on consumer surplus measures.

The Introduction gives an extensive but summary review of the literature (40 references), referring to a more detailed review by Cherchi and Polak (2004, not available for review). It finds that the treatment of the impact on welfare of taste heterogeneity has been rare, particularly outside the transport field (but see von Haefen, etc., on recreational evaluations), while in the transport field most authors ignore income effects as being unnecessary.

Section 2 reviews the relevant aspects of welfare economics, introducing (Marshallian) consumer surplus (CS) and compensating variation (CV), pointing out that if calculations are to be made by integrating the demand curve it is usually necessary to use CS, at least as an approximation for CV. CV might be more attractive, because it includes substitution and income effects, whereas the CS only gives substitution effects (see chapter 2). It is stated that CV is difficult to calculate and assumptions are usually therefore made in practice to make it tractable.

The first such assumption is that ‘at least one of’ the disturbances is GEV – equations 3, 4 etc. however assume that *all* of the disturbances take that form. With this assumption, the integral of the demand curve is the convenient ‘closed’ (generalised) logsum form. But, the authors claim, only if indirect utility is linear in income can the expected CV be calculated in this way, because otherwise the averaging is affected by the non-linearity. Most people would find it simpler to say that the logsum gives the CS and this is only equal to the CV if income ‘falls out’ of the model, i.e. appears linearly.

The second assumption is then that indirect utility is indeed linear in income. This does not solve the above tractability problem, because the authors allow the possibility that residual money has a different value as a function of the alternative chosen, so that a third assumption that the linear function is the same for all alternatives is required to derive the standard (generalised) logsum result for CV (= CS). It is noted that a mixed GEV model can be used, simply applying the mixing distribution (e.g. by sampling) to derive a mixed logsum.

The fourth assumption is that a GEV model without mixing can be used (although we could note that this is not directly relevant to CV/CS calculation but more to demand

model estimation and application). The constancy of the money coefficient that this assumption implies means that the logsum can be translated to money terms (note that the important assumption concerns the money coefficient *only*).

Finally, the rule-of-a-half approximation is presented, giving the CV as the sum over all OD pairs and modes of $\Delta CV = -\frac{1}{2} \Delta GC D_{tot}$, where ΔGC is the change in generalised cost and D_{tot} is the total of the demand before and after the change (somewhat strangely described)¹⁷. It is noted that this formula does not depend on the precise form of the demand function.

Section 3 presents some simulation experiments which look at the impact on the logsum of model specification errors. A simulated data set was constructed which included both taste and income variation. Exact benefit measures were calculated.

In Section 4, these exact measures were compared with values calculated from models which ignored the full heterogeneity. The conclusions in Section 5 indicate serious bias in these simplified measures, which is particularly due to taste variation (most notably in the marginal value of money). In practice, this problem may be concealed by the cancelling of opposite errors. The authors are reluctant to generalise but warn against using simple measures in complex situations.

3.19 **Daly (2004)**

The objective of this paper is to improve the understanding of the GEV modelling framework.

The first section describes the foundations of the utility maximisation paradigm, which is based on the assumptions of completeness, transitivity and either continuity or finiteness. Behaviour can then be described ‘as if’ a real-valued quantity, which can be called utility, was being maximised. The RUM approach is then adopted to allow for observation limitations on the part of the observer or random variation (e.g. mistakes) on the part of the consumer. Choice is then a random event. This approach is consistent with the classical RUM literature.

The second section and the associated Appendices show that *all* the ‘classical’ RUM literature, i.e. up to and including the work of Börsch-Supan (1990, which is essentially an investigation of failure of the first assumption) is based on three further assumptions concerning the distribution of the random term:

1. it is continuous and proper;
2. it is bounded; Technically this means that the absolute value of utility is greater than M gets very small, if M is big enough; In practice it means that alternatives

¹⁷ The simplest formulation is perhaps $\Delta CV = -\Delta GC D_{avg}$, where D_{avg} is the average of demand before and after the change. This is the same as $\Delta CV = -\Delta GC D_{before} - \frac{1}{2} \Delta GC (D_{after} - D_{before})$, which is more recognisable formulation, showing the rule-of-a-half in its basic form.

are never unavailable ($U = \text{minus infinity}$) or that people are captive ($U = \text{plus infinity}$);

3. it does not depend on any component of the non-random utility.

The last assumption is variously described as translational invariance, defining location parameter models, etc., and essentially eliminates taste variation. Many models satisfy these criteria.

In the third section, Daly goes on to show that the three assumptions above correspond one-to-one with the requirements for McFadden's GEV theorem. It then follows that every model satisfying the three invariant RUM assumptions also satisfies the GEV theorem, i.e. has a welfare function as derived by McFadden, which for models based on GEV distributions is a logsum or generalised logsum. All models based on GEV distributions are invariant RUM.¹⁸

In the fourth section, Daly introduces the concept of budget constraints, arguing that money budgets are different in kind from time budgets, because spending time inevitably involves some notion of intrinsic 'enjoyment' ('enjoyment' might have a negative connotation). The crucial difference is that time has an inevitable hedonistic component.

The fifth section discusses the impact of income effect in terms of the Slutsky equation and its consequences for invariant RUM models. It is shown that it is perfectly possible for models to satisfy the GEV theorem, i.e. to have a McFadden welfare function, without requiring the AIRUM assumption, as discussed in McFadden's later work.

The final main section discusses appraisal measures, where Daly attempts to draw together the recent literature (as reviewed here). A device is presented for simplifying the McFadden sampling procedure, which is possible providing there is no satiation with respect to income.

In conclusion, it is noted that the applicability of the GEV theorem is much wider than is generally thought, so that functions derived from it can be used generally, although the problem of aggregation remains. Models incorporating taste variation can also be used but here there may be issues to resolve (Daly was not aware of the work of von Haefen).

¹⁸ Thus the class of models satisfying the GEV theorem is larger than the class of invariant RUM models, which in turn is larger than the class of models based on GEV distributions.

CHAPTER 4 **Review of papers on the application of the logsum as an evaluation measure**

4.1 **EXPEDITE consortium (2002)**

General idea

The starting point for the EXPEDITE project was the question how in forecasting and policy simulation at the European scale one can benefit from the detailed knowledge on transport behaviour and reactions to policy measures embodied in a number of existing national models. In other words: how can one extend this knowledge to a study area comprising the current member states of the EU, Switzerland and Norway?

Within EXPEDITE, five national models were available for passenger transport (in the order in which they were originally developed):

- the Dutch National Model System (NMS or LMS);
- the Norwegian National Model (NTM-4);
- the Italian National Model (SISD);
- the Danish National Model;
- the Swedish National Model (SAMPERS).

Integrating the outcomes of the underlying models for passenger transport

In the first part of the EXPEDITE project, a large number of runs have been carried out (up to 80 runs per model) with each of the above mentioned national models. To the maximum possible extent, the same runs were done with each of the models. For the base-year (1995), outcomes were generated in the form of 'levels matrices'. The levels matrices for tours give the number of tours per person per year by mode and distance band. A 'tour' is defined as a round trip, starting and ending at home. The levels matrices for passenger kilometres give the number of kilometres travelled per person per year, by mode and distance band.

The modes are:

- car-driver;
- car-passenger;
- train;
- bus/tram/metro;
- non-motorised.

Distance band is the other dimension of the levels matrices. Seven bands are used, ranging from 0 to 160 km (one-way distance).

There are different levels matrices (tours and kilometres) for five travel purposes and for many population segments. The socio-economic and demographic population segmentation used in the meta-model is as follows:

- age distribution (<18, 18-<65, >=65);
- gender (male, female);
- occupation of persons (employed; not employed);
- household size (1, 2, 3, 4+);
- household net income class (0-11300, 11300-18200, 18200-29500, 29500-38600, 38600 Euro per year);
- car ownership (person in a household without a car, person without a driving licence in a car-owning household, person with a licence in a household that has more driving licences than cars, person with a licence in a household that has at least as many cars as licences).

Besides levels matrices for 1995, the outcomes of the national model runs also consist of switching matrices: changes in tours or in passenger kilometres (same units as the levels matrices), as a result of a change in a policy-related model input variable. There are switching matrices for changes in the running cost of the car, travel times by car, and for cost, in-vehicle time, wait and transfer time and access/egress time of train and bus/tram/metro. Runs for different percentage changes (e.g. +10%, + 25%, +40%, -10%, -30%) were carried out, because the travel demand response to cost and time changes may very well not be linear.

For each segment, the levels and switching matrices in tours and kilometres from all five national models were averaged (unweighted) to get the 'prototypical' matrices that are used in the meta-model to forecast for Europe.

The zoning system in the meta-model is the NUTS2 level. At this levels there are around 250 zones. For each zone, expansion factors were calculated depending on the importance of the population segments in the zone (many of these weights could be zero for a specific zone). By multiplying the tours and passenger kilometres from the prototypical matrices with the expansion factors, initial predictions for each of the zones are derived. These are forecasts for all travel demand generated in the zone, with one-way distances up to 160 km, by mode, distance class, travel purpose and population segment. The trip distances

over 160 km are missing here, because several national models used have only a limited coverage of long distance travel.

These initial forecasts are first corrected for differences in travel behaviour by area type and by road and rail network type, based on runs with the Dutch national model, the ANTONIN model for the Paris region and the SCENES model for Europe. The model forecasts for 1995 that result after applying the area and network type correction factors have been validated against observed data on the use of each mode (if available by distance class), by country. This has resulted in a set of mode-specific, distance-class-specific and country-specific correction factors, which are also kept in forecasting. In this way, the meta-model accounts for 'residual' factors affecting travel demand, such as climate, hilliness and historical developments.

This meta-model for passenger transport also includes area-wide speed-flow curves to take account of the feedback effect of changes in congestion due to policies that change the amount of car use.

To obtain forecasts for all distance bands (the meta-model for passenger transport, based on the national models is for travel up 160 km), results from the SCENES model for travel above 160 km can be added to those of the meta-model.

Calculating the impact of policy bundles

For a change in travel time or cost for which the national models have not been run (e.g. +20%), we could have derived the switching matrices by linear interpolation between the matrices of a 10% change and a 25% change. This would amount to assuming a piece-wise linear response to cost changes. However, the meta-model accounts for the non-linearities in the response to policy changes by going back to the original logit formulation, as used in the national models. This method is also used to calculate the impact of a policy bundle.

A policy bundle is a combination of individual policy measures (e.g. increase in car cost and decrease in public transport cost). A limited number of policy bundles have been tested in the national models, and change matrices for these bundles are directly available for use in the meta-model. For all other policy bundles, the meta-model calculates the effects of the combination of policy measures from the results of individual policy measures, taking account of non-linear effects in the following way.

- sub-additivity: the combined effect is less than the sum of the separate effects
- super-additivity: the combined effect is more than the sum of the separate effects.

The method used can lead to both types of effects, depending on the location on the logit curve. As an example we study the combined effect of an increase in the car running cost of 25% and a decrease in the train and bus/tram/metro cost by 25%.

The switching matrices for both of these measures in isolation are available from the national model runs.

The model now calculates probability matrices P_{mdp} (m indicates mode, d distance class and p population segment) by dividing all numbers in the levels matrix of a segment by the total in the bottom-right cell for:

- the levels matrices T_{mdp} ;
- the levels matrices with policy 1: $T_{mdp} + D^1_{mdp}$;
- the levels matrices with policy 2: $T_{mdp} + D^2_{mdp}$.

It is further assumed that:

- the non-linearities in the responses of the meta-model to policy measures are due to the logit nature of the underlying utility-based models;
- the average utility of the shortest distance band for the non-motorised modes will remain unchanged in any forecast scenario (this is just a standardisation).

Now the average utilities (standardised by the utility of the shortest distance band for the non-motorised modes) can be calculated from the probability matrices as follows.

The general formula for the multinomial logit model is:

$$P_{mdp} = \frac{e^{U_{mdp}}}{\sum e^{U_{mdp}}}$$

Therefore:

$$\ln(P_{mdp}) = U_{mdp} - \ln(\sum e^{U_{mdp}})$$

and:

$$U_{mdp} = \ln(P_{mdp}) + \ln(\sum e^{U_{mdp}})$$

The same can be done for the average utility of the shortest distance band for the non-motorised mode. The standardised utility for mode m, distance class d and primitive p then becomes:

$$\text{Standardised } U_{mdp} = \ln(P_{mdp}) - \ln(P_{m=\text{non-motorised}, d=\text{shortest}, p})$$

Given that the starting point are the 'p's, i.e. e^U / e^U , a scale standardisation is needed to recover comparable logsums $\ln(e^U)$ as between base and forecast/scenario. Since the forecasts/scenarios did not have pedestrian schemes or bike lanes, short distance non-motorised was chosen as base.

What is effectively done here is that the underlying utility functions are calculated: the utility functions that are not specific to any one of the national models, but apply to EXPEDITE as a whole.

Now one can calculate these average utility matrices for each of the three situations (base, base with policy 1, base with policy 2). Then to obtain the utility matrix of the policy bundle 1&2, the utility of the base is added to the utility change of policy 1 and the utility change of policy 2. After that the outcome is standardised by using the utility of the shortest distance band for non-motorised as the base. The result is transformed to probabilities (by exponentiation). The resulting probability matrices for the base with the policy bundle 1&2 are below.

Measure of welfare change

These underlying utility functions are also used to calculate the change in the logsum, that is caused by a policy measure or bundle.

$$\text{Logsum} = \ln(\sum e^{U_{mdp}})$$

This gives the change in consumer surplus, and can be segmented by population segment to analyse how different population segments are affected by a policy. twenty-one different policies (and a reference 2020 situation) have been tested using the EXPEDITE meta-model. The logsum difference between the policy run and the reference run for 2020 was calculated. This difference needs to be attributed a monetary value. In a simple linear additive discrete choice model without income effects, this could be done by using the cost coefficient (by travel purpose). However, the EXPEDITE model does not contain cost coefficients (by travel purpose), but uses results from underlying national model runs (as described above), which contain different cost coefficients (sometimes in logarithmic form). As an approximation a car cost increase by 10% for the year 1995 (which is an amount of money that is known) has been used to establish conversion factors (implied average cost coefficients) for the logsum, with a distinction per travel purpose.

4.2 **Odeck, Rekdal and Hamre (2003)**

Introduction

This paper investigates the magnitude of socio-economic benefits that may be gained by converting the current road financing oriented cordon toll in Oslo to a road congestion-pricing scheme. It is hypothesized that such a move will generate considerable socio-economic benefits as compared to status quo. A transport model with four alternative congestion-pricing regimes is developed to assess the hypothesis.

Theoretical approach logsums

For estimation of the socio-economic benefits from transport policies, the following expression has been used:

$$\text{Socio-Economic Benefit} = \text{Consumer surplus} + \text{Revenues from congestion pricing} + \text{Income from public transport} - \text{Operating costs for public transport.}$$

The consumer surplus change includes both time and direct money changes (operating costs, parking costs, fares, tolls).

To estimate the user benefits, an often applied method is “the rule of the half”: a change in consumer’s surplus is estimated as changes in generalized costs times the average demand before and after the introduction of the project:

$$CS = \frac{1}{2} (GC_1 - GC_0)(X_1 - X_0)$$

where GC_0 = generalized costs before implementation; GC_1 = generalized costs after implementation; X_0 = demand before implementation; X_1 = demand after implementation.

However, because there are changes in generalized costs both for motorists and public transport users in applying congestion pricing and also because in the rush hours models¹⁹ the changes also vary with departure time, this paper uses an alternative method of estimating the consumer surplus, based on the property of the logit model. Changes in the logsum of the logit model are conceptually equivalent to the traditional consumer surplus (rule of half method). The logsum can be expressed as:

$$LS = \frac{\ln(\sum_{it} \exp(GC_{it}))}{\beta_c}$$

where GC_{it} = is the generalized costs expression in the model for mode i at departure time t . β_c = the model parameter for monetary costs.

4.3 Castiglione, Freedman and Davidson (2003)

Introduction

Activity-based models are often seen as attractive alternatives to traditional trip-based travel demand forecasting models. New policy analysis requirements demand that forecasting models represent travel choices and the geographical, temporal and behavioural context in which these choices are made. Activity-based models can incorporate these details. This paper describes the activity-based San Francisco Model and its application in the evaluation of the proposed New Central Subway project downtown San Francisco. The application involves the policy analysis of the change in consumer surplus through the use of the United States Federal Transit Administration's (FTA) "User Benefit" methodology and software.

The San Francisco Model

The San Francisco model is a tour-based micro-simulation model, which forecasts daily activity patterns for individual San Francisco residents. This model has been used in transportation planning since 2000. The model employs the full daily activity pattern approach within a disaggregate micro-simulation framework. The main feature of the full day pattern approach is that it is simultaneously predicting the main components of all of a person's travel across the day. A synthesized population of San Francisco residents is input to the component models of vehicle availability, day pattern choice (tour generation), tour time of day choice, destination choice and mode choice. Both destination and mode choice are predicted at both the tour and trip level. The synthesized tours and trips are aggregated to represent flows between traffic analysis zones before traffic assignment. The model predicts the choices for a full, representative sample of residents of San Francisco County, almost 800,000 simulated individual person-days of travel.

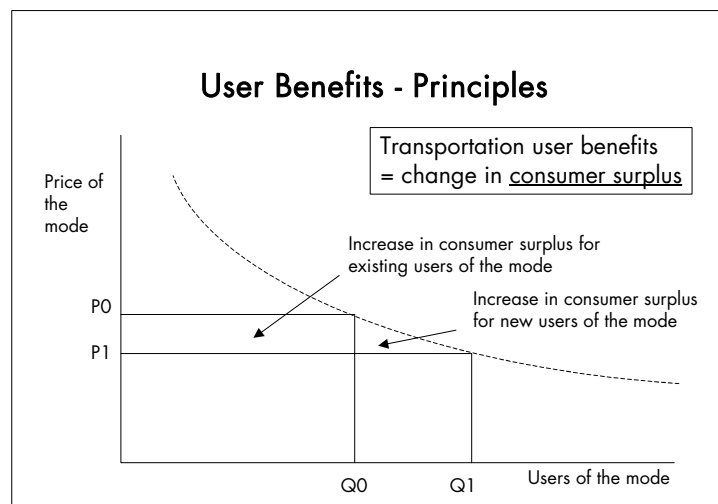
¹⁹ A distinction has been made between morning rush hour and afternoon rush hour and the models representing the rush traffic periods split the traffic into departure hours (for each of the three hours of the rush periods)

A micro-simulation framework, using logit modelling, is applied to individuals and households making vehicle ownership, trip pattern, and trip destination and mode choices. A Monte Carlo method is used to select outcomes according to these logit model probabilities based on random number draws. Each time the sequence of random numbers used to simulate choices is varied, the model result (“end state of the model”) may change.

Theoretical Approach to logsum

The FTA’s “User Benefit” methodology intends to qualify the benefits resulting from a project accruing to all transportation system users by using a common unit of travel-time savings. This measure supports two project justification factors: mobility improvements and cost effectiveness. The measure is based upon basic economic theory, measuring the change in consumer surplus attributable to a new transportation investment (Figure 3).

Figure 3. User benefits principles



The measure is based on two key inputs. The first is derived from the denominator of the logit mode choice model, the so called logsum. This is a measure of the composite utility of all modes of travel considered by the model. An increase in the mobility attributable to the introduction of transportation infrastructure will increase the utility of that mode and will be reflected as an increase in the composite utility of all modes. Dividing this by the in-vehicle time coefficient will convert the logsum to a composite price of travel for all modes, measured in equivalent minutes of in-vehicle time.

The second key input is the incremental costs of the project, defined as the estimated annualized capital cost (not including the financing costs) plus annual operating and maintenance costs. The transportation system user benefit is defined as the incremental annual cost of the project divided by all annual travel-related benefits in terms of hours saved by all (existing and new) users of the system.

Calculation of this measure is supported through software distributed by FTA called SUMMIT, according to the following formulas:

1) Multinomial logit mode choice with two alternatives, car and transit:

$$P_{trn} = \frac{e^{U_{trn}}}{e^{U_{trn}} + e^{U_{auto}}}$$

2) Natural log of the denominator of formula 1 is the mode choice logsum

$$e^{U_{trn}} = \frac{P_{trn}}{1 - P_{trn}} \times e^{U_{auto}}$$

3) User benefits are measured by zone-pair for nine market segments (current maximum SUMMIT) by multiplying the difference in price of travel

$$Cost_{(auto+transit)} = \frac{\ln(e^{U_{auto}} + e^{U_{transit}})}{\beta_{ivt}}$$

4) Difference between build alternative and base alternative price measures consumer surplus. Results can be aggregated to show total user benefits or user benefits by zone or district aggregation

$$Surplus_{ijs} = T_{ijs} (C_{ijs(base)} - C_{ijs(build)})$$

4.4 Ecorys Transport and 4cast, May 2004

Introduction

This report describes the possibilities to improve the translation of model output of the National Transport Model (LMS) and the New Regional Model (NRM) in travel time benefits (for Cost-Benefit analysis). It distinguishes between short term improvements and long term improvements and only briefly mentions the use of logsums as a long term improvement that may have both advantages and disadvantages

Theoretical approach logsums

Logsums are defined as the mean utility of a set of possible choices and it is a good measure for example to determine the accessibility from a zone. The change in mode-, destination- and time-choice for all destinations and modes are incorporated in the logsum from the LMS. The logsum is defined as the natural logarithm of the sum of the exponential expected utility for all choice alternatives.

In traditional cost-benefit analysis, only changes in utility due to changes in travel times and costs are taken into account. Time is converted into monetary costs using values of time. Using logsums has the advantage that all changes in utility are accounted for, i.e. due to any change in any variable that figures in the utility function (for example parking availability or comfort variables).

To translate logsums into money, a method would be to translate the change in utility into time-units and combine this with a mean or representative Value of Time to transfer it to costs.

Practical considerations using logsums

For research studies that are not using LMS/NRM but other types of transport models, this method cannot be used to compare results.

Another consequence is that not all the levels of detail that via the “old” method could be calculated can be calculated using the logsum method without additional research and/or adaptations to the model. Examples are: distinction of the utility per trip-link, distinction of the utility per time of day period, distinction of the utility per mode type.

4.5 **Gupta, Kalmanje and Kockelman (2004)**

Introduction

This work takes a look at the impacts of new toll roads, as well as possible bridge and downtown cordon toll policies at Austin, Texas. An integrated transportation-land use model, based on a rather standard four-step Travel Demand Model (TDM) process and a model for distribution of households and employment (DRAM-EMPAL model), was applied to the Austin region in order to anticipate the traffic, land use and welfare impacts of various pricing policies. Feedback within the TDM model (from traffic assignment back into trip distribution, mode and time of day choice models) was performed using a method of successive travel cost averages. Results from the TDM then were used in the land use models to predict household and employment locations under various tolling scenarios. Resulting changes in travel demand were evaluated to appreciate the TDM-predicted changes in traffic patterns, destination choices, mode and time of day decisions, locational accessibility, network level of service, toll revenues, land use patterns, and travel-based measures of welfare.

Theoretical approach logsums

Estimates of daily travel-related benefits from road pricing alternatives were estimated as consumer surplus, at the destination choice level for the average person residing in each zone. The approach is based on the method explained in the Kalmanje and Kockelman [2004] paper (see below).

Consumer surplus is the difference in the maximum expected utility of one’s destination choice opportunities before and after a change in the travel environment, as shown in this equation:

$$CS_{i,p} = \frac{1}{\alpha_p} (E(Max(V_{i,p}))^{ps} - E(Max(V_{i,p}))^{np})$$

where α_p is the marginal utility of money (specific to each trip purpose) and is the product of the estimated coefficients on cost (in the mode-departure time model) and generalized cost or logsum (in the destination choice model), ps and np denote the pricing and no pricing (status quo) scenarios, and

$$E(Max(V_{i,p})) = \ln \left(\sum_{j \in C} e^{V_{i,j,p}} \right)$$

where $V_{i,j,p}$ denotes the utility of person at origin i choosing destination j for trip purpose p , with C denoting the full choice set of all possible destinations.

In this work, this measure of consumer surplus is not applicable for Home-Based Work (HBW) trips since the destination choice is fixed. Instead, the following equation is used to compute consumer surplus for HBW trips²⁰. It is the difference in expected maximum utility levels derived across all modes and departure times for a particular destination, and is multiplied by the probability of choosing that destination ($P(j)$).

$$CS_{i,p} = \sum_{j \in C} \frac{P(j)}{\beta_c} (LOGSUM_{i,j,p}^{ps} - LOGSUM_{i,j,p}^{np})$$

where $LOGSUM_{i,j,p}$ is the generalized cost between an origin-destination pair (i,j) and is defined as the negative of the maximum expected utility derived across all mode and departure time combinations for a trip purpose p :

$$LOGSUM_{i,j,p} = -\ln \left(\sum_{m,t} e^{\beta_{t,p} Time_{i,j} + \beta_{c,p} Cost_{i,j} + \beta_{m,t,p}} \right)$$

where β_t , β_c and $\beta_{m,t}$ are the coefficients on time, cost and the alternative-specific constants in the joint mode-departure time choice model.

Average daily consumer surplus is calculated for an individual residing in zone i by aggregating consumer surplus for home-based trips using the average daily number of trips per individual.

Practical considerations of this approach

The assumptions used in the TDM process impose certain behavioural limitations. For example, trip generation is assumed to be inelastic with respect to travel costs and thus unchanged following the introduction of road pricing policies. In reality, some trips may be wholly suppressed, while others emerge (due to latent demand) under the effects of pricing. The static traffic assignment procedure and the choice models employed deal only with homogeneous users limiting predictions of policy impacts to the average user level rather than across user groups. The travel time and travel cost skims from traffic assignment are computed based on generalized cost. The procedure may be skimming for shortest path using toll roads if they satisfy the generalized cost criterion. This is not a serious limitation; but, in the presence of heterogeneous users, this may bias the welfare computations slightly.

²⁰ The Austin-calibrated TDM models rely on four trip purposes: home-based work (HBW), home-based non-work (HBNW), non-home-based work (NHBW), and non-home-based-non-work (NHBWNW)

4.6 Kalmanje and Kockelman (2004)

Introduction

This paper explores the possible transport and property value impacts of a new congestion management policy called credit-based congestion pricing (CBCP). Congestion Pricing (CP) is based on the principle that if people pay the true marginal cost of road use, the congestion externality is internalised and lower but more efficient levels of road use result during peak periods. This approach has many advantages, but a major drawback is the adverse equity impacts on low-income groups and others with special travel needs. To tackle this equity issue, the authors propose a revenue-neutral policy (CBCP), where tolls generated from marginal cost-pricing are returned to all licensed drivers in a uniform fashion as a sort of driving “allowance”.

Using destination, mode and departure time choice models sensitive to changes in travel times and costs, household travel demands were simulated in order to appreciate the transport effects of a CBCP policy for Austin, Texas. Changes in housing values as a result of CBCP were also simulated. The trip-based welfare impacts of such a policy were compared for three scenarios (full network pricing, major highway pricing only, and no pricing), in order to identify households and neighbourhoods that will benefit most and least from such policies. The results corroborate prior results and hypotheses about the potential of a CBCP policy to alleviate congestion and generate benefits across the region and traveller types.

Theoretical approach logsums

Consumer surplus at the destination choice level after the policy change has been used as the welfare measure in this paper. This provides a comprehensive measure of impact, capturing impacts on destination, mode and departure time of day choices. The following equation gives the expression for consumer surplus under a policy change; it is the difference in the expected maximum utilities before and after the policy change.

$$CS_{i,p} = \frac{1}{\alpha_p} (E(Max(U_j)_{ip})^n - E(Max(U_j)_{ip})^0)$$

where n and o denotes the new and old policies and α_p is the trip purpose specific destination choice model's marginal utility of money, in this case equal to the multiple of the coefficients on cost (in the mode-departure time model) and generalized cost or logsum (in the destination choice model).

This measure is based on a given origin, assumed to be one's neighbourhood of residence, and it is computed as a logsum of utilities of all destination choices from that origin with $V_{i,j,p}$ denoting the utility of person at origin i choosing destination j for trip purpose p , with C denoting the full choice set of all possible destinations²¹:

²¹ The (systematic) utility of a destination from a particular origin is given by:

$$V_{i,j,p} = \beta_{i,p} \text{LOGSUM}_{i,j,p} + \beta_{s,p} \ln(\text{SIZE}_{j,p}) + \ln(\text{ATTR}_{j,p})$$

$$E(Max(U_j)_{ip}) = \ln \left(\sum_{j \in C} e^{V_{i,j,p}} \right)$$

Consumer surplus ($CS_{i,p}$) under CBCP has been calculated for an average individual located in every zone i for each trip purpose p using the corresponding destination and mode-departure time choice models. The rebate per trip (R) was calculated by dividing the daily allowance by an individual's daily average number of trips. Average daily trip making of 4.6 trips/individual²² was assumed to calculate the average CBCP allowance for each trip, and a trip-weighted average of the trip-based welfare measures produced the final measure (CS) shown in the following equation:

$$CS = \frac{\sum_p (CS_{i,p} + R) * \#_p}{\sum_p \#_p}$$

where $\#_p$ is the average number of trips by each purpose per individual.

Results

The results of the model application to various policy scenarios suggest that most Austin residents would be better off under policies that employ CBCP, whereas relatively few would benefit under a policy of congestion pricing, without beneficial revenue redistribution.

Results of the welfare analysis for both the CBCP policies tested (system-wide CBCP versus CBCP on main roads only) show that the potential “winners” under these two policies can be very different. While users near the central business district (CBD) are predicted to gain from pricing only the main roads, users living in northwest Austin (close to several major employers) gain the most under system-wide CBCP.

Since tolls charged are based on a higher Value of Travel Time (VOTT) estimate than those obtained from the model estimates, welfare changes under CBCP for an average Austin resident may be largely (or exclusively) positive throughout the region.

where $SIZE_j$ is either the area or the employment at the destination zone “j”, $ATTR_j$ is the attraction at the destination zone “j”, and $LOGSUM_{i,j}$ is the “logsum” for the origin-destination pair (i, j), computed as:

$$LOGSUM_{i,j,p} = -\ln \left(\sum_{m,t} e^{\beta_{t,p} Time_{i,j} + \beta_{c,p} Cost_{i,j} + \beta_{m,t,p}} \right)$$

The logsum is the negative of the expected maximum utility derived across all mode and departure time combinations for that particular destination, and it is a measure of accessibility of that destination from origin of interest (i), where t , c and m,t are the joint mode-departure time model coefficients (on time, cost and the alternative-specific constants, respectively).

²² Data derived from the 1998-1999 Austin (household) Travel Survey (ATS)

4.7 Koopmans and Kroes (2004)

In this article the costs of queues in the whole of the Netherlands are calculated using logsums as provided by the Dutch National Transport Model (LMS) for the year 2000. The exact way in which the logsums are computed is not specified in the article, but the standard LMS software (including NSES) has been used for this.

The costs of queues are determined as follows:

1. Perform an LMS run for the year 2000 using free-flow travel times for cars, and determine the consumer surplus using the standard logsums output, expressed in travel time equivalent
2. Do the same thing but then using congested travel times at equilibrium
3. Determine the difference between the logsums obtained in 1. and 2.
4. Multiply the results of 3. by a value-of-time measure; in this case an average value-of-time of 9 Euro per hour has been used. The marginal utility of money was not used here to convert the outcomes into money units, because in the LMS costs appear in a logarithmic form, and because the focus in the study on the amount of time lost (in hours).
5. Gross-up the average weekday results obtained in 4. by the number of working days per year to obtain a yearly total
6. Apply some corrections to account for international traffic, freight traffic and the observed differences in total numbers of cars between NSES and the observed base matrix.

The results indicate that for the year 2000 the total congestion cost on the highway network amounts to 1.5 Billion Euro. This is substantially higher than the estimated costs using the more traditional method based upon “vehicle-hours lost” (VHL) as obtained in the assignment: that method obtained an estimate of 0.8 Billion Euro for 2000. This happens to be the same number that was estimated for the year 1997 using measurements of queues and queue-lengths by AVV.

The authors hypothesise, based upon a simple linear demand model, that for increasing levels of congestion the ratio of the estimated costs by the logsum method relative to the VHL method will further increase.

The relevance of this paper is that it shows, for the first time, what the outcomes of the (changes in the) logsum calculation applied using the LMS are. It can be concluded that the order of magnitude of the results is plausible, and that the results are higher than those obtained using the more traditional VHL method using the same model. This is conform what the authors expected.

4.8 De Raad (2004)

The main research question of this MSc Thesis of TUD student Paul de Raad was:

How do outcomes of the logsum method relate to traditional congestion costs computations based upon travel time losses?

Secondary research questions included:

- *What is the difference between the logsum method and conventional computations when analysing the costs and benefits of future infrastructure projects?*
- *How do outcomes of both methods relate for different levels of congestion?*
- *What is the effect of the use of the logsum method on the feasibility of infrastructure investment?*

In the Thesis two types of examples have been elaborated:

1. A test simulation of a substantial capacity increase for the A13 between Rotterdam and Den Haag using the Dutch national Transport Model (LMS). The results indicated that the traditional method estimated the annual benefits at 0.8 Million Euro per year, while the logsum method arrives at 3.1 Million Euro per year. So the benefits estimated using the logsum method are almost a factor four higher than those using the traditional Vehicle Hours Lost (VHL) measure using the same model.
2. A test to see how the results vary for different levels of congestion; here OD matrices for three years (1986, 2000 and 2020) have been assigned to a road network for the year 2000. The results were as follows:

Method	Matrix 1986 Network 2000	Matrix 2000 Network 2000	Matrix 2020 Network 2000
Traditional VHL	0.3 Billion Euro/year	0.7 Billion/year	1.9 Billion Euro/year
Logsum method	0.9 Billion Euro/year	1.3 Billion/year	2.0 Billion Euro/year

Both methods show increases in congestion costs with increasing level of congestion. But the VHL results appear to increase more rapidly than the logsum results, which is different from what was expected.

In this test the results of the logsum method were also used to illustrate how this method can provide interesting indications of where the impact of congestion is most severe: instead of indicating where the queues are located the method indicated which households living in which zones suffer most from the congestion.

The main interest of this Thesis is that it illustrated how cost/benefit calculations based upon the logsum method may lead to different conclusions concerning infrastructure development, given that higher benefits are typically computed for these infrastructures. Another interesting but unexpected result was the fact that the logsum results appeared to

increase less rapidly with increasing congestion level than the traditional VHL method. And finally the illustration of the spatial impact of traffic congestion was potentially interesting.

4.9 **RAND Europe (2004)**

The aim of this pre-study was to analyse the influence of accessibility (represented by means of logsums and travel times) on settlement behaviour of residents and firms (jobs). This first part of the research has identified realistic changes in the accessibility variables of the TIGRIS XL land-use model by analysing existing LMS (Dutch National Transport Model). Existing government policies with respect to infrastructure development and land-use indicate that a change of about 2.5% for the logsum variables and about 10% for the travel times are realistic values.

More specifically the impact of four different factors on accessibility (measured using logsums and travel times) has been assessed:

1. The expected changes in accessibility as caused by autonomous developments and existing policies
2. The influence of infrastructure developments
3. The influence of congestion charging and transfer from fixed cost to variable cost
4. The influence of alternative spatial developments (including labour following residential locations).

Accessibility has been observed to change through time as a consequence of increases in income, changes in the composition of the population, and possible policy measures. All these developments affect accessibility. However, it is not possible to say *a-priori* what the impact will be on the logsums or travel times. The anticipated development between 2010 and 2020 has been seen to lead to counteracting effects between increased income and infrastructure on the one hand, and different spatial patterns and increased population on the other hand.

Providing additional road infrastructure improves both the logsums and the travel times, but the effect diminishes as distances to the new infrastructure projects increase.

Congestion charging and variabilisation are examples of policy measures that have been simulated. Here the impact is negative for the logsum measure, but positive for the travel time.

The spatial development has a more local impact: around new residential areas the logsum values increase due to jobs following residential locations. But travel times increase due to increased congestion.

Accessibility measures appear to be most sensitive to pricing policies and infrastructural changes. Spatial developments have little impact at the national level, but the impact can be important at local level. Changes of accessibility over time are a mix of the above mentioned effects, with possibly counteracting effects. Travel time is a more sensitive

measure for accessibility than the logsum: the observed changes are significantly larger. The logsum measure is sensitive to changes in travel time, but also to changes in cost, employment and population, which may have an effect that works against travel time.

The table below summarises the observed ranges of the changes in accessibility measures for different zones:

Accessibility measure	Change over time (2010-2020)	Infrastructure development	Pricing policies	Spatial development policies
Logsum (move/stay)	0.0-5.0%	0.0-1.0%	0.0-3.0%	0.0-0.5%
Logsum (household type A)	0.0-10.0%	0.0-0.5%	0.0-2.0%	0.0-0.5%
Logsum (household type D)	0.0-4.0%	0.0-1.0%	0.0-3.0%	0.0-0.5%
Travel time	0.0-4.0%	0.0-7.0%	0.0-5.0%	0.0-0.5%
Logsum "business"	0.0-2.0%	0.0-0.5%	0.0-0.5%	0.0-0.5%
Logsum "commuting"	0.0-4.0%	0.0-0.5%	0.0-2.0%	0.0-0.5%
Freight travel time	0.0-3.0%	0.0-6.0%	0.0-3.0%	0.0-0.5%

The relevance of this paper is that it shows, for the first time, what the sizes are of the changes in logsums that can be expected for different factors in forecasting. Important observations are also that:

- Changes in logsums tend to be smaller than changes in travel times;
- The direction of changes in logsums may be the same as those in travel times (e.g. due to infrastructure projects) but also the opposite (e.g. road pricing).

4.10 **US Department of Transportation, Federal Highway Administration (2004?)**

Introduction

This case study compared the use of a regional travel demand model, SACMET96, with two transport-land use models, MEPLAN and TRANUS, for testing regional transportation and land use policies. The two separate modelling approaches that are used are:

- Using a travel demand model only. Under this approach, alternative transport investments and policies are simulated by changing network characteristics in the travel demand model. The effects of land use policies are simulated by changing the geographic distribution of future development. This approach is used widely by metropolitan areas throughout the United States.
- Using a travel demand-land use model. Under this approach, an integrated model is used that includes feedback between transport, accessibility and land

development. It also allows land use policies, such as development fees or zoning changes, to be modelled directly by influencing the price and availability of land.

The project has evaluated a range of policies, both individually and in combination. Impacts were measured for the years 2005 and 2015 on travel, emissions, user benefits, and the spatial distribution of population and employment.

Theoretical approach logsums

In addition to travel, emissions, and land use changes, the study also measured overall user benefits. The total monetary value of cost savings, time savings, and other benefits were measured based on the difference in the overall "utility" of travel between two alternatives. Capital and operating costs for each alternative are then subtracted, to obtain the net change in benefits. User benefits are calculated for both personal and commercial vehicle travel. Overall benefits are reported on a cost-per-trip basis.

This case study uses the approach of Small and Rosen (1981) to measure consumer benefit, also known as compensating variation (CV). This measure can be obtained from discrete choice models:

$$CV = 1/\lambda \left\{ \left[\ln \sum_{m \in M} e^{V_m(p^f)} \right] - \left[\ln \sum_{m \in M} e^{V_m(p^0)} \right] \right\} \quad (1)$$

where λ is the individual's marginal utility of income, V_m is the individual's indirect utility of all m choices, p^0 indicates the initial point (before the policy change), and p^f indicates the final point (after the policy change). The change in indirect utility is converted to dollars by the factor $1/\lambda$, or the inverse of the individual's marginal utility of income.

The compensating variation formula as given above is adapted to work with the mode choice models used in this case study (SACMET96). In these models, households are segmented into income/worker categories and person trips are generated for those categories. To obtain compensating variation for each income/worker category h , the following formula is applied for all modes m and for all trips Q between all origins i and all destinations j :

$$CV_h = 1/\lambda_h \left\{ \sum_{i=I} \sum_{j=J} \left[\left(\ln \sum_{m \in M} e^{V_{mh}(p^f)} * Q_{ijh} \right) - \left(\ln \sum_{m \in M} e^{V_{mh}(p^0)} * Q_{ijh} \right) \right] \right\} \quad (2)$$

where λ is provided by the coefficient of the cost variable in the mode choice equations. Total compensating variation is obtained by summing the compensating variation obtained from each income/worker group. Compensating variation is also obtained from the non-home-based mode choice models, which are not stratified by household/income classes.

The specific steps followed to calculate the compensating variation are as follows:

1. Extract from the travel model (in this case, MinUTP) the logsum (denominator) of the mode choice equation for each trip type and household/income category, as well as the number of trips in each category.

2. Use this to calculate CV for each category. Summing this value would provide the total user benefits, without including project costs.
3. Determine the estimated annualized capital and operating costs of each alternative.
4. Determine the proportion of Vehicle Miles of Travel (VMT) attributable to each trip category, based on number of trips and average trip length by type.
5. Allocate the annualized capital and operating costs to trip category based on the proportion of VMT in that trip category.
6. For each trip category, subtract the allocated costs from the CV to produce the net economic benefit. Summing across trip types gives the net benefits by income group.

This method is applied for personal vehicle travel, for commercial vehicle travel the economic benefits are calculated based simply on travel time and operating costs savings.

Practical considerations of logsum approach

Consumer benefits were calculated from the mode choice models SACMET96, but not from the transportation land-use model MEPLAN. There are differences in opinion over whether consumer surplus measures such as the Small-Rosen traveller welfare model described above, are valid when land use demand shifts. At this moment, research takes place to test methods obtaining measures of consumer welfare from similar transportation-land use models (Dr J. Hunt, University of Calgary using the Portland statewide model, which is a TRANUS-application).

CHAPTER 5 Outcomes from the LMS case study

5.1 The 'classic' and the 'improved classic' method

Four different runs were carried out with the LMS:

- The existing LMS for the reference situation 2020.
- The existing LMS for the project situation 2020 (same as reference, except for the implementation of Rondje Randstad, with particular speed and frequency increases for a number of train links between the big cities in the Randstad and reductions on some of the minor train links).
- An alternative (simplified) specification of the LMS, without nested logit (NL) models (for specific travel purposes these were replaced by multinomial logit (MNL) models), for the reference situation 2020.
- An alternative (simplified) specification of the LMS, without nested logit models (for specific travel purposes these were replaced by multinomial logit models) for the project situation 2020 (same as reference, except for the implementation of Rondje Randstad, with particular speed and frequency increases for a number of train links between the big cities in the Randstad and reductions on some of the minor train links).

The latter two runs were included to analyse the differences in the evaluation results (especially for the logsum results) that are caused by changing the model specification. In this way, we performed a sensitivity analysis to check for the sensitivity to differences in model specification. For several travel purposes the LMS mode-destination choice models are nested logit models (where there is for instance for some travel purposes more substitution between destinations than between modes). The alternative specification for which we calculated logsums was derived by estimating all mode-destination choice models as multinomial logit models (i.e. with cross-elasticities for changes in for instance time and cost that are the same for all choice alternatives). We did not just fix the tree coefficients in the nested logit models to be equal to one (the formal requirement for reduction to multinomial logit), but re-estimated the models as multinomial logit models, where the other coefficients could change as well. Some mode-destination choice models in the LMS already were multinomial logit models, and were not changed. The mode-destination

choice model for the travel purpose 'other' could not be estimated as MNL and remained a nested logit model.

Below are the results for the number of tours train travellers from the first two runs (using the LMS as it exists today and is regularly used in applications). Please note that for HB (home-based) business and 'other' travel, the number of train tours is predicted to decrease. This is due to the fact that the project variant does not provide train times that are better than the reference situation 2020 in all cases; for some origin-destination relations, the train times in the reference are shorter. The travel times by train in the project variant are always at least as good as in the base year, but the reference situation²³ also includes some improvements in train travel times compared to the base year, some of which (especially stop trains) are not in the Rondje Randstad variant.

Table 1. Number of tours by train on an average working day in 2020 according to the current LMS

	Reference 2020	Project 2020	Project 2020 (Reference=100)
commuting	505571	513088	101.5
HB-business	16659	16509	99.1
NHB-business	18959	18978	100.1
education	170121	171955	101.1
shopping	37499	37584	100.2
other	112678	112455	99.8
total	861487	870569	101.1

In the 'classic' method, the benefits from the project are calculated as follows. For instance for commuters the number of travellers that stay in the train is taken to be 505571. The number of 'new' travellers is taken to be $513088 - 505571 = 7571$ (in fact this is substitution from other modes). The average travel time (train in-vehicle-time) in the reference situation is 62.26 minutes for commuting. With the project this is 61.62 minutes. This time gain is used as the benefit for all stayers: $505571 * (62.26 - 61.62) * (\text{value of time for train commuters})$. This value of time comes from the 1997/1998 stated preference (SP) surveys that Hague Consulting Group carried out for AVV. For the new travellers the gain is (rule of half): $7571 * (62.26 - 61.62) * (\text{value of time for train commuters}) * 0.5$. Repeating this for all purposes gives the traveller benefits as in Table 2 (first row).

As can be seen from the above example and the numbers of train travellers in Table 1, the benefits to travellers are completely dominated by the benefits of those that stayed in the

²³ The reference situation run was taken from Ecorys and 4cast (2004).

train. The new train passengers only make up about 1% of the total amount of train passengers in the project situation. Furthermore the benefit for a new train passenger is calculated as only half that of a remaining train passenger.

In the last column of Table 2 are outcomes when including not only the train in-vehicle-time benefits, but also the gains in terms of in-vehicle-time (bus) during the access to the train station and during the egress from the train station (this is exogenous project input). This more than doubles the time benefits of the project. The increase is caused by the fact that the (large) railway stations that will be used more in the Rondje Randstad variant have better bus/tram/metro connections, so the access and egress times will be shortened.

The benefits in Table 2 are between 0.10 and 0.23 Euro per train tour.

We did not calculate additional benefits for the car users due to the reduction of congestion on the roads (that would be caused by substitution from road to rail), because the substitution from road to rail was so small (it's about a third of the 'new' train travel) that the average travel time by road did not change. The results in Table 2 can not be compared to those for Rondje Randstad in Ecorys and 4Cast (2004), because in this report from 2004, the differences between the project situation and the reference situation do not only include the Rondje Randstad project itself but also changes to the zonal distributions (different configuration of cities).

Table 2. Project benefits to travellers using 'classic' methods for the full year 2020 in millions of 2003 Euros²⁴

	Train in vehicle time only	Including access/egress time
Current LMS	24	58
LMS with MNL only	24	58

The travel purposes for children up to twelve year old (education and other travel) and freight traffic are not taken into account, both in the calculations with 'classic' methods and in calculations with the logsum method.

The outcomes in numbers of tours for the simplified version of the LMS (no nested logit models) are given in Table 3. The traveller benefits are in the bottom row of Table 2.

²⁴ The calculations were originally made in 1997 Euros (the year of the SP value of time surveys). For business travel a factor 1.23 was used to go from 1997 to 2003 (average contractual wage rate increase) and 1.30 for the increase in the values of time between 2003 and the forecast year 2010 (Ecorys and 4Cast, 2004). For commuting these values were 1.18 (consumer prices) and 1.15, and for other travel the values were 1.23 and 1.11. For the conversion from an average working day to a full year we used the factor 285 (Ecorys and 4Cast, 2004).

Table 3. Number of tours by train on an average working day in 2020 according to the simplified (re-estimated) LMS

	Reference 2020	Project 2020	Project 2020 (Reference=100)
commuting	505164	512624	101.5
HB-business	16654	16545	99.1
NHB-business	18723	18789	100.1
education	170106	171942	101.1
shopping	42165	42638	100.2
other	112678	112455	99.8
Total	865490	874992	101.1

The demand forecasts from the MNL specification are very similar to those for the specification that is actually used in project evaluation, except for the shopping tours. The difference is caused by the changes in the alternative-specific constants in the shopping model, which are much larger than for other purposes. However for the difference between project and reference situation this does not matter.

The benefits for the travellers from these time gains are in the second row of Table 2. At the given level of precision, the NL and MNL specifications give the same benefits (with greater precision there are differences between NL and MNL).

In Ecorys and 4Cast (2004) a number of short-term improvements to the ‘classic’ method is proposed and discussed, leading to what we call here ‘the improved classic method’. In the remainder of this section we shall go through these proposals one by one and discuss the Rondje Randstad project in the light of these short-term improvements.

Q-hours versus average speeds.

In project evaluations in the recent past, travel time gains on the road networks have sometimes been restricted to changes in vehicle-hours lost due to congestion (Q-hours). This is not correct, since changes in the speed of non-congested traffic produce welfare just as well. All travel time gains should be included in the project assessment. For the Rondje Randstad project this improvement has no consequences, because the substitution from road to rail is so small that the Q-hours and the average travel times by road do not significantly change.

Modal shift: values of time of new mode or average over old and new modes

The ‘classic’ method uses the rule-of-half and calculates the benefits for new travellers (with some (?) mode) as half of the benefit for travellers that stay with that mode. At a workshop on the direct effects it was proposed to use the average of the values of time of

the old and new mode here. This is also proposed in Ecorys and 4Cast (2004). Later on, this was criticised for not being consistent with the underlying welfare economics (and the rule of half in particular). We agree with this criticism: the short-term ‘improvement’ is not actually an improvement, and the ‘classic’ method should be used instead. That is what we did earlier in this section for the Rondje Randstad project.

Including and weighting different travel time components of the journey by public transport

This improvement refers to public transport journeys and is relevant for the evaluation of the Rondje Randstad project. The ‘classic’ method is then to use the gains in the in-vehicle-times for the main mode only. This is what we did above for train by only using the train in-vehicle-times. The improvement consists either of including the access, egress and transfer times without differential weighting or including these components with a weight per minute (based on the literature) that is different from that of in-vehicle-time. Above we also applied the first approach of implementing the improvement, by including the access/egress time changes. This leads to substantial increases in the benefits. In Ecorys and 4Cast (2004), the inclusion of other time components also led to additional benefits (but relatively smaller; this must be due to the fact that the definition of the project also includes spatial changes). The weighting further increased those benefits by 25%. In the logsum calculations, all time gains are included, and weighted according to the coefficients of the LMS utility functions (no weights from the literature).

Distinguishing income groups

Another proposed improvement is to calculate the benefits by income group. For the total benefit to travellers this does not really matter (in Ecorys and 4Cast, 2004, the difference in total benefits to travellers was only 2%), but the information about the distribution of gains/losses over income groups can be valuable in its own right. The above ‘classic’ calculations did not distinguish between income groups. In the logsum calculations we did include the distinction by income class.

Impact of changes in comfort and punctuality

The ‘classic’ approach does not include improvements in comfort and punctuality. In Ecorys and 4Cast (2004) some assumptions (necessarily somewhat *ad hoc*) were made on how the project might improve comfort (especially being able to sit in the trains instead of having to stand) and punctuality, and how improvements in comfort and punctuality would be valued. This resulted in an increase of the travel time benefits of 6% for the comfort improvements and 2% for the punctuality improvements. The calculations in this report, both for the ‘classic’ and for the logsum approach, do not include comfort and reliability benefits. Such effects could be included in the LMS runs that provide the demand forecasts as well as in the LMS logsums provided we would have estimates on:

- The improvement in comfort and reliability that results from the project evaluated. For reliability of travel times, the new LMS-BT module

(Kouwenhoven, van Grol en Kroes, 2004) can provide estimates of changes in reliability.

- A conversion of (unit or percentage) change in comfort or reliability to variables that are in the LMS utility functions (such as travel time, or mode-specific constants); Some of this is available from De Jong et al. (2004) and RAND Europe (2005).

Distinguishing time of day periods

In Ecorys and 4Cast the impacts are also given by time of day period (morning peak, evening peak and off-peak). For the train this led to 8% more time benefits and 2% less for car. In the explanation, it is argued that the differences with the 'classic' approach depend on the generated travel. Since in our Rondje Randstad calculations we have very limited generated traffic (about 1% of the total travel in the project situation; and this constitutes modal shift), this distinction will not influence the results. Shifts between time periods can in principle be captured by a logsum approach, but the LMS version 7 does not have a time of day choice model for train travellers (the LMS version 8 will have), but only for car trips.

Chain mobility and related transport policy ('flankerend beleid')

These were also in the list of short-term improvements, but were not included in the LMS runs in Ecorys and 4Cast (2004) for the 'improved classic' method, and have not been included in this report either.

5.2 The logsum method

Logsums (in utils) were first calculated for the reference situation 2020 and for the project situation with shorter train times because of the 'Rondje Randstad' project. This project is the only difference between the project situation and the reference situation.

In total we have four sets of logsums

- For the reference situation, using nested logit models (NL).
- For the project situation, using nested logit models.
- For the reference situation, using multinomial logit models (MNL).
- For the project situation, using multinomial logit models.

Logsum differences for the difference between the reference and the project situation were calculated for the nested logit models and for the multinomial logit models.

The logsums and logsum differences were originally calculated per tour. Then these outcomes were aggregated/expanded to logsums and logsum differences for combinations of travel purpose and income class (with five income classes, as used in the LMS).

For each of these two logsum types of differences (NL, MNL), we applied two different methods for the conversion from utils to money:

- Method 1: Translate from utils to minutes, using the LMS travel time coefficients (by purpose) and then translate from minutes to 1995 guilders by using the recommended values of time (from the 1997/1998 stated preference surveys that Hague Consulting Group carried out for AVV) by purpose and income group. Because the project studied (Rondje Randstad) is a rail project, and rail users are affected most, we used the time coefficients (for in vehicle time and other time components) of rail here. The values of time used are those by income class and travel purpose (not by mode, but over all modes).
- Method 2: Divide the logsums in utils by the product of the LMS cost coefficients and the expected value of $1/\text{costs}$ per tour (all by income class and purpose). The costs enter the calculation because of the use of logarithmic costs in the LMS mode-destination choice models. In a linear cost model, division by the cost coefficients would have been sufficient for conversion to money units. Moreover, the use of the expectation of $1/\text{costs}$ is only approximately correct. On the other hand, this method does not use the information on values of time from the SP survey and therefore does not suffer from inconsistency of using one set of implied values of time in the LMS to get the transport demand impacts and another set of values of time from the SP surveys for conversion into money. Method 1 does have this consistency problem, and so has the present use of external values of time in OEI. The cost coefficients in the LMS are the same for all modes (differ only between purposes), but a problem is the treatment of modes and population groups with zero costs (slow modes, car passengers, students). The LMS itself uses zero if there are no cost and $\ln(\text{cost})$ for positive cost. For our conversion to money in method 2, we need to divide by costs, and have to avoid division by zero. To calculate this, we used the lowest observed cost (we found that this is just below 1 guilder, and used 1 guilder here) per tour for modes and groups with zero costs, so that these have will have a small impact on the final results.²⁵

Moreover, we also applied a benefit calculation method that does not require monetisation of utils:

- Method 3: According to the Startnotitie (AVV, 2004), Nol Verster of Ecorys suggested to do a simulation in which the prices in the model are raised to a level where the logsum in the project situation would be equal to the value in the reference situation. The total extra amount of money paid by the population would then be the monetary benefit for the travellers of the project. In Ecorys and 4cast (2004), this is described as the cost increase that would bring the behavioural effect back to zero. We interpreted this as: calculate the amount of money (transport costs in the LMS) needed to bring the utility back to the level in reference situation (the compensating cost change). This is then no longer strictly speaking a logsum measure, but a measure of the compensating variation.

²⁵ A cost formulation of the form $\ln(\text{cost}+1)$ in the LMS would have been more convenient. This also gives zero when cost is zero and the proper derivative for zero costs.

In all three methods, the travel purposes lower education ('basisonderwijs') and 'other trips by children under twelve' were excluded. For freight trips, no utility functions were available and no logsums were calculated.

Below the three methods are described formally (for a given purpose and person type):

We have utility functions of the form:

$$U = \alpha_c \ln[C] + \alpha_t T + \dots$$

In which:

C: cost in 1995 guilders

T: time in minutes

α_c : 1 guilder is α_c utils, or 1 util is $1/\alpha_c$ guilders

α_t : 1 minute is α_t utils, or 1 util is $1/\alpha_t$ minutes

We also define:

LS = logsum in utils

Now method 1 and 2 work as follows:

Method 1 (Value of time):

$$LS/\alpha_t = \text{logsum in minutes}$$

$$\text{Logsum in guilders}^{26} = (LS/\alpha_t) \cdot \text{VoT}$$

The VoT comes from an external model, estimated on stated preference data.

Method 2 (Average cost):

Conversion from utils to money:

$$E(\partial U / \partial C) = E(\alpha_c / C)$$

Therefore we get:

$$E(\partial U / \partial C) = \alpha_c \cdot E(1/C)$$

$$\text{Logsum in guilders}^{27} = LS / (\alpha_c \cdot E(1/C))$$

²⁶ When the model is a two-level nested logit (such as the LMS mode-destination choice models for most purposes), the time coefficient needs to be multiplied by the logsum coefficient (the differential of the logsum with respect to time is $\alpha_t \cdot \gamma$, where γ is the logsum or tree coefficient that needs to be between 0 and 1 for global consistency with utility maximisation).

²⁷ When the model is a two-level nested logit (such as the LMS mode-destination choice models for most purposes), the cost coefficient needs to be multiplied by the logsum coefficient (the differential of the logsum with respect to time is $\alpha_c \cdot \gamma$, where γ is the logsum or tree coefficient that needs to be between 0 and 1 for global consistency with utility maximisation).

Method 3 (method ‘Verster’)

This method calculates the benefits of a new infrastructural project by calculating how much the costs in a utility function must rise to maintain the same level of total utility when time has decreased (as a consequence of the new project).

In our case the new project is a train LOS improvement, so costs and time reflect the mode ‘train’.

Suppose we know the train cost and train time for an average person in an income class i with purpose p . Then, for each average person, the utilities for the reference and project run should be equal. In formula (subscripts for income class and purpose are left out for readability):

$$\begin{aligned}
 U_{ref} &= U_{var} \\
 \alpha_t T_{ref} + \alpha_c \ln C_{ref} &= \alpha_t T_{var} + \alpha_c \ln C_{var} \\
 \alpha_c \ln C_{var} &= \alpha_t (T_{ref} - T_{var}) + \alpha_c \ln C_{ref} \\
 C_{var} &= e^{\frac{\alpha_t (T_{ref} - T_{var}) + \alpha_c \ln C_{ref}}{\alpha_c}}
 \end{aligned}$$

The project makes the train travel time in the project variant T_{var} smaller than in the reference T_{ref} . C_{var} is a hypothetical cost level in the project variant that brings the utility back to the reference level. The increase in train costs $C_{var} - C_{ref}$ to maintain the same utility value reflects the benefit of the new train project. Strictly speaking, the application of this method is not really a logsum approach (the logsum changes in utils are not used), and more similar to the compensating variation (see chapter 2, but it can be questioned whether the specification of income in the LMS is sufficient to pick up the income effect as in Hicksian demand functions and the related compensating variation). It is essentially the same method as proposed by McFadden (1996) and developed by Daly (2004), as described in chapter 3 of this report. Unlike the logsum calculation, a transfer from utils to money is not necessary.

In this implementation of the Verster method, we only calculate the benefit for tours that use the train in both the reference and project situation. Only the train utility functions are used. We do not include benefits for other modes and do not analyse substitution here. The reason for excluding substitution in this method is that for substitution it would lead to inconsistent behaviour. Some tours might shift from road to rail and have a longer travel time in the project situation (rail) than in the reference situation (road). For these persons we would calculate a time loss and a negative compensating cost. But this is inconsistent with the fact that transport level-of-service has improved and substitution takes place within a utility maximisation framework. A solution to this would be to use the rule-of-half for those who substitute (which gives the right sign since it only looks at the travel

time improvement for rail), but that would mean going back to the ‘classic’ methods. Therefore we decided not to include substitution in this method.

The outcomes for the total travellers’ benefit change for the year 2020 are in Table 4.

Table 4. Traveller benefits (project minus reference) for the full year 2020 in millions of 2003 Euros²⁸

Method:	LMS nested logit models	LMS multinomial logit models
Logsum using SP values of time	44-51	4-6
Logsum using average costs	56	42
Compensating variation (Verster method)	134	146

Differences between the first and second method are due to the fact that the first method uses external values of time (not from the LMS), whereas the second only uses information from the LMS. If values of time as implied by the LMS coefficients would have been used, both methods would have produced approximately the same total monetary change. Generally speaking the SP values of time are larger than the implied average LMS values of time for commuting for the lowest income classes. Also the SP values of time exceed the LMS values for business travel, shopping and other purposes. For commuting for the higher income classes (these are important categories for train travel), the SP values of time are lower than those implied by the LMS.

For the current LMS specification (with nested logits), the outcomes are in the same range for the first two methods of monetising the logsum. For method 1 with the current LMS we have different results depending on whether we do the calculations by income group (first value given) or for all incomes at the same time (second value given). The benefits calculated using the Verster method are clearly higher, even though these only include the benefits to the current train travellers and not the substitution effect (which is probably small in this case, given the limited amount of substitution). By definition the compensating variation (CV) will not be equal to the consumer surplus (CS), and the first two rows give CS measures, while the third row gives a CV measure. The theory says that for a price reduction the CV will exceed the CS (because of the income effect). That is what we find here, but the magnitude of the difference is larger than could be expected from an income effect, and is probably due to only using coefficients from the train utility functions instead of all utility functions as in methods 1 and 2. For the calculation of a

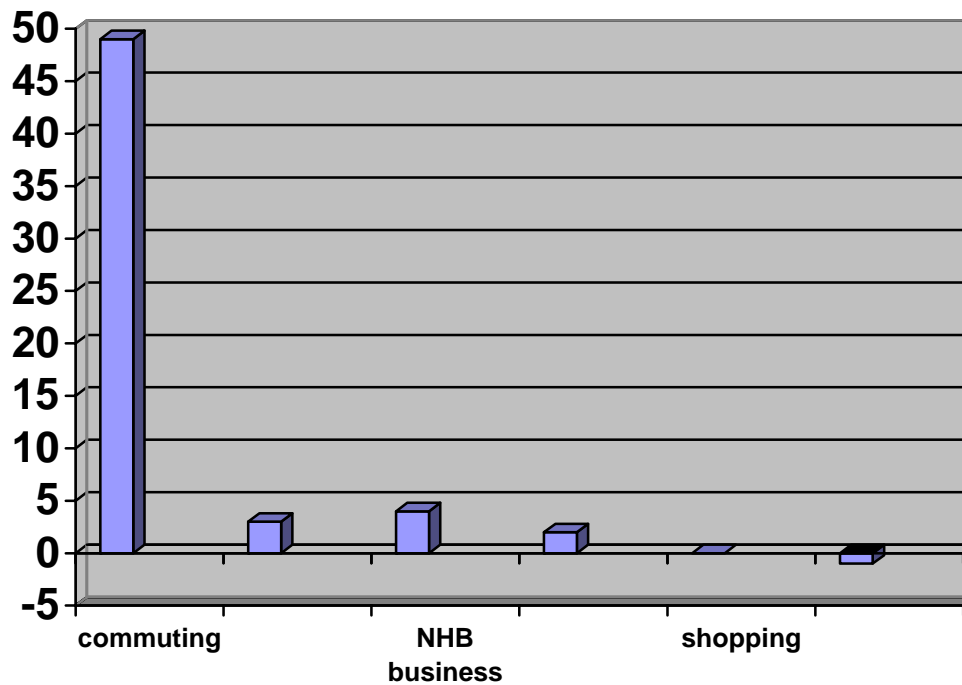
²⁸ The calculations were originally made in 1995 guilders (as used in the current LMS version). For business travel a factor 1.28 was used to go from 1995 to 2003 (average contractual wage rate increase) and 1.30 for the increase in the values of time between 2003 and the forecast year 2010 (Ecorys and 4Cast, 2004). For commuting these values were 1.23 (consumer prices) and 1.15, and for other travel the values were 1.23 and 1.11. For the conversion from an average working day to a full year we used the factor 285 (Ecorys and 4Cast, 2004).

proper CV (e.g. section 3.13, 3.15 and 3.19) a change in income (or in the costs of all modes) would be required to bring us back to the utility of the reference situation. A change in the costs of only one mode will need to be larger than this, which is what we can see in the outcomes. Furthermore, the logsum calculations in method 1 and 2 are based on specific origin-destination pairs, including pairs with reductions in train time and pairs with increases. Method 3 only uses the average train travel time reduction (per income group and purpose).

If instead of the current LMS we would use a multinomial logit specification we get very different results for first method. The outcomes of applying method 1 for nested logit and MNL are quite similar for the non-business purposes, the difference is caused by business travel. What happens here is that the MNL gives a notable decrease in the logsum value as a result of the project, especially for business travel. In the first method, this reduction in utils gets a large weight, since the external value of time for business travel is much higher than the values of time for all other purposes and higher than what is implied by the LMS itself. For method 2 (with lower implied values of time for business travel) and the Verster method, the differences between multinomial logit and nested logit are rather small. The latter method does not use the logsum change per purpose in utils, but calculates the compensation in cost terms that would be required to compensate the travel time reduction for those who stay in the train and does not lead to negative outcomes for some of the purposes.

The project benefits for travellers (according to method 2 and the nested logit models) by travel purpose are in Figure 4.

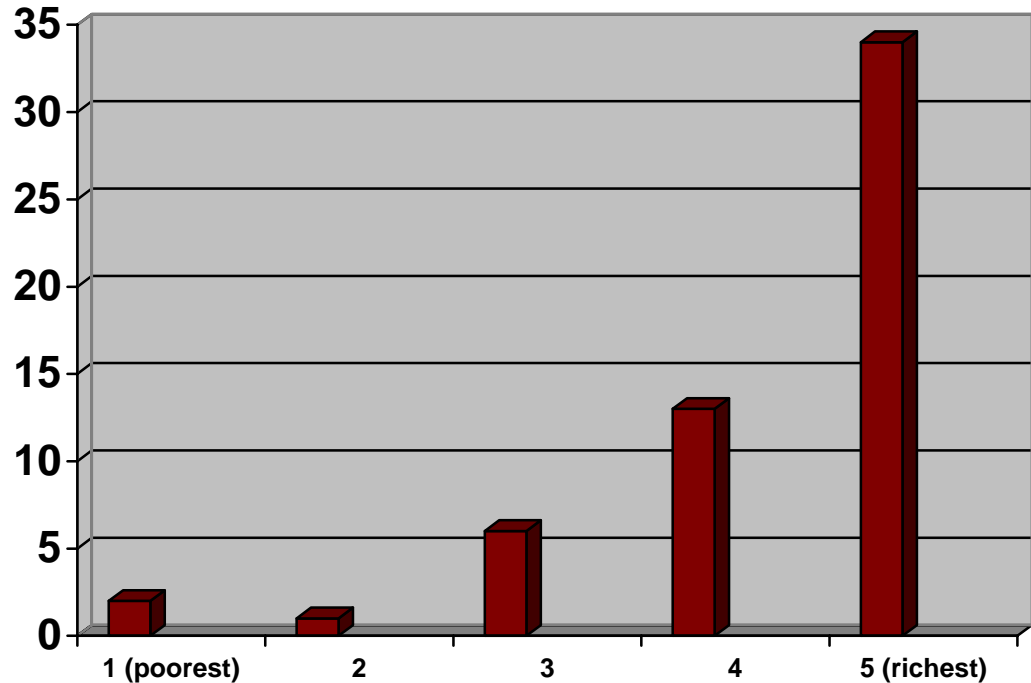
Figure 4. Division of logsum change for Rondje Randstad over purposes (in millions of 2003 Euros for the full year 2020)



It is clear that most of the benefits of Rondje Randstad are accruing to commuters, with business travellers coming in second and persons going to school/university third. The division over purposes is similar when using method 1 or 3 (for the current LMS specification). One should keep in mind that about 58% of the train tours and 55% of the train kilometres (on an average working day) in the LMS are made by commuters.

Furthermore, the higher incomes groups are enjoying most of the benefits, as can be seen from Figure 5. Again the figure is for the second method and nested logit, but for both method 1 and 3 the distribution over income groups is very similar (for the current LMS specification). The share of the highest income group in total train tours in the LMS is 55% and it is also 55% in total train kilometres.

Figure 5. Division of logsum change for Rondje Randstad over income groups²⁹ (in thousands of 2003 Euros for the full year 2020)



So the outcomes of the monetised logsum change (especially method 1) turn out to be very sensitive to model specification (shift from nested no multinomial logit). But we would like to emphasise that for the LMS as it is used in practice, this difference does not occur and the outcomes of two different monetarisation methods of the logsum gave rather similar outcomes for the effect of the project on travellers.

5.3 Comparing results of 'classic' and 'logsum' methods

The logsum outcomes for the nested logit models for the first two monetarisation methods are similar (see Table 5) to those using the in train in-vehicle time and access/egress time gains and SP values of time ('classic' method). We think this is just a coincidence. The logsums take into account the changes in all components of the utility functions: in-vehicle time, access/egress times, but also wait time. In the logsum approach, the LMS time coefficients are used to go from gains in utils to gains in minutes. These coefficients are not consistent with the SP results used in the 'classic' method. Most of the values of time from the SP are higher than those of the LMS (at the average costs): see Table 6 below.

²⁹ The income bands for net annual household income are, after conversion to Euros: 0-11300; 11300-18200; 18200-29500; 29500-38600; >38600.

Substantial differences can be found especially for shopping, other purposes and for business (the LMS value of time does not include the employers value of time for business travel, as the SP value of time does).

Table 5. Traveller benefits (project minus reference) for the full year 2020 in millions of 2003 Euros³⁰

Method:	LMS nested logit models
Classic method, train in-vehicle time only	24
Classic method, including access/egress time	58
Logsum using SP values of time	44-51
Logsum using average costs	56
Compensating variation (Verster method)	134

Table 6. Ratio of SP value of time to average implied LMS value of time

Income category purpose	1	2	3	4	5
Commuting	1.835	1.205	0.957	0.640	0.993
HB Business	2.359	1.393	1.657	1.579	2.560
NHB business	1.317	1.474	1.708	1.675	2.924
Education	1.027	0.525	1.068	0.962	1.479
Shopping	16.790	15.605	16.079	15.751	21.324
Other	2.529	2.479	2.660	2.647	3.640

³⁰ The calculations were originally made in 1995 guilders (as used in the current LMS version). For business travel a factor 1.28 was used to go from 1995 to 2003 (average contractual wage rate increase) and 1.30 for the increase in the values of time between 2003 and the forecast year 2010 (Ecorys and 4Cast, 2004). For commuting these values were 1.23 (consumer prices) and 1.15, and for other travel the values were 1.23 and 1.11. For the conversion from an average working day to a full year we used the factor 285 (Ecorys and 4Cast, 2004).

CHAPTER 6 **Summary, conclusions and recommendations**

6.1 **Summary and conclusions**

The theory on the use of the logsum change as a measure of the change in the consumer surplus, to be used in project appraisal, was published in the late seventies and early eighties. Nevertheless, the application of this theory in practical appraisals of transport projects has been very limited, and most applications in transport evaluation that the authors are aware of have been undertaken only recently (after 2000). It is not easy to find the reasons for the inertia to use the theory in applied work. To some extent it can be related to the complexity of some of the theoretical literature, but the basic logsum concept (with constant marginal utility of money) is fairly straightforward to apply. It may also have to do with the fact that in many countries there is no (national) model system based on disaggregate random utility models. For the computation of logsum changes, disaggregate Generalised Extreme Value (GEV) models, such as the multinomial logit and the nested logit, are required, although in the EXPEDITE project it proved possible to go back from a more aggregate model to the implied underlying utility models. National disaggregate transport models are in use in Scandinavia, the Netherlands and Italy and regional and urban models using these concepts can be found in the same countries, France, the United Kingdom, Australia, Israel and especially the United States. It is therefore not surprising that the logsums applications in evaluation took place in the USA, Scandinavia and The Netherlands. It is unlikely that the computer run times for the calculation have been a major obstacle for the use of logsums in evaluation, since all the required inputs are already computed in the standard procedures for application of disaggregate models (calculation of individual probabilities in sample enumeration).

The applications that were reviewed in this report are summarised in Table 7. All applications use models that include mode choice. Some logsum applications in project evaluation also use models for destination choice and/or departure time choice. Logsums are first calculated for each individual decision-maker in the sample, and then aggregated over groups of decision-makers. Various segmentations are used, also depending on the segmentation used for the time or cost coefficients used later on to convert utils to money or time. A common segmentation for the logsum calculations and outputs is by travel purpose. The applications of the logsum concept in transport project appraisal all use the

Table 7. Summary of applications of the logsum in transport project appraisal

Model application	Choices included	Segmentation	Marginal utility of income	Conversion method
San Francisco (Castiglione et al, 2003)	Mode choice	By zone pair and 9 segments based on household size and car availability	constant	Using a common in-vehicle time coefficient to get outcomes in minutes
Europe (EXPEDITE, 2002)	Mode-destination choice	By almost 1,000 person segments and 5 travel purposes	constant	Using an implied cost coefficient per purpose to get outcomes in euros
Austin (Gupta et al., 2004; Kalmanje and Kockelman, 2004)	Mode, destination and departure time choice	4 trip purpose, calculated for an individual resident	constant	Using a cost coefficient per purpose to get outcomes in dollars
The Netherlands LMS (Koopmans and Kroes, 2004; De Raad, 2004)	Mode, destination and departure time choice	8 travel purposes	constant	Using time coefficients per purpose to get minutes, then using value of time
Oslo (Odeck et al, 2003)	Mode-departure time choice	By trip purpose	constant	Using a cost coefficient per purpose to get outcomes in Kroner
The Netherlands TIGRIS (RAND Europe, 2004)	Mode, destination and departure time choice	8 travel purposes	constant	No conversion used, calculation of changes in utils
Sacramento (USDoT, 2004)	Mode choice	Household segments base on income/worker categories	constant	Using a cost coefficient per segment to get outcomes in dollars

relatively simple formulation with constant marginal utility of money. It could be dangerous to assume that the marginal utility of income would be constant over a wide income range (it is more likely that it will decline with increasing income). Theory has moved beyond that in the nineties, but the later formulations are not in practical use³¹.

Two applications (Castiglione et al, 2003 and Koopmans and Kroes, 2004) do not convert the logsum change in utils directly into money units, but convert to time in minutes. The other applications use one or more cost coefficients to get outcomes in money units.

A case study in the logsums as an evaluation measure was carried out with the Dutch National Model System (LMS), for a rail project in the Randstad area ('Rondje Randstad'). We applied two different ways of monetarisation of the logsum change in utils: method 1 that uses SP values of time and method 2 that uses average costs. We also implemented a compensating variation method (method 3 or Verster method).

We found that the application of the conventional rule-of-half approach leads to very different results depending on whether only the train in-vehicle time changes or also the access/egress time changes are taken into account. The logsum results for this project also vary, between different monetarisation methods and even more between different model specifications (nested versus multinomial logit). But for the specification as used in practice (nested logit), the differences in outcomes for two different ways of monetising the logsums are rather small. Most of the project benefits accrue to commuters and the highest income group, who make more than half of the train tours and kilometres.

6.2 Recommendations

At present the method used in The Netherlands for quantifying the benefit for travellers of a transport project consists of calculating the change in consumer surplus (in terms of a reduction of generalised travel costs) for both the current users of the directly affected alternative and for new users. For the latter group the rule-of-half is used. This procedure has a basis in welfare analysis. For projects at a national scale, the LMS is often used to produce demand changes and the resulting benefits in travel time and costs. For regional projects, the NRM (new regional models) are regularly used, which use similar demand functions as the LMS.

We think that replacing this approach by the logsum approach would provide a number of advantages:

³¹ However, for policy analysis not the absolute levels but only changes in consumer surplus are relevant, and the simple equation for calculating the expected consumer surplus can be used if the marginal utility of income is constant over the range of implicit changes that are considered by the policy. So, for policy changes that change the consumer surplus by small amounts per person relative to their income, the formula can be used even though the marginal utility of income in reality varies with income (Train, 2003). Also please note that the Sacramento application uses the assumption of a constant marginal utility of money only within a number of distinct income/worker categories. This provides a solution, which we also used for the LMS applications where cost coefficients also differ between five income groups.

- When using logit models as in the LMS, the logsum change also gives the change in the consumer surplus, and in a more exact way than the rule-of-half does, which is based on a linearisation.
- At present there is an inconsistency in the evaluation procedure: for calculating the changes in travel demand. The LMS is used, which has its own set of implied values of time. Then the resulting time changes are monetised using a different set of values of time (from Stated Preference surveys, SP). When using logsums we can avoid the use of external values of time (except in a method, that we called 'method 1', of monetisation that expresses the logsum change in minutes first, and then through SP values of time in money). On the other hand, the SP studies might contain information that the LMS is lacking and it would be even better to estimate the transport demand models on a combination of the available Revealed Preference (RP) and SP data.
- The logsum method might seem to be much more complicated than the rule-of-half, but in fact a major advantage of logsums is the ease of calculation. Particularly when several alternatives are changing, e.g. in a destination and time period choice when traffic is reassigned in response to a project, the rule-of-half calculations can get very complicated while the logsum ones are easy and need to be done anyway to get demand. The logsum method can also easily give results per population group (the conventional approach can do this as well, but this is often a lot of extra work).

An advantage of the conventional approach is that it is more transparent (but only in simple situations) and more intuitive and therefore easier to explain to non-experts. On the other hand, the transport models that produce the logsums are already common practice.

We think the advantages of the logsum approach outweigh the disadvantages, but we also think that further testing of the logsum method is required. We recommend to test the logsum approach for other schemes, especially for highway schemes (or a combined road and public transport scheme) where the purpose and income mix is likely to be more representative of the travelling population as a whole and where the substitution effects might be more important. The monetisation of logsum changes through external values of time (as in method 1) does not seem attractive, because that would bring the consistency problem back in. Therefore at this stage we would prefer to use the monetisation of the logsum with the average costs (method 2). The compensating variation method (method 3) avoids using the average costs to approximate the effect of logarithmic costs, but can only give the effect for tours that do not change mode, which is not an attractive option either.

References

Please note: Papers and reports for which a review can be found in this report are listed in **bold** characters.

Adviesdienst Verkeer en Vervoer (2004) Startnotitie Case studie Logsum benadering, AVV, Rotterdam.

Bates, J. (2003) Economic evaluation and transport modelling: theory and practice, resource paper for 10th International Conference on Travel Behaviour Research, Lucerne.

Ben-Akiva, M. and Lerman, S. R. (1979) Disaggregate travel and mobility choice models and measures of accessibility, in Hensher, D. A. and Stopher, P. R. (eds.) *Behavioural Travel Modelling*, Croom Helm.

Ben-Akiva, M. and Lerman, S.R. (1985) Discrete choice analysis: *Theory and application to travel demand*, MIT Press, Cambridge Massachusetts.

Börsch-Supan, A. (1990) On the compatibility of nested logit models with utility maximisation, *Journal of Econometrics*, Vol. 43, pp. 373-388.

Castiglione, J., Freedman, J. and Davidson W. (2003) Application of a tour-based microsimulation model to a major transit investment, San Francisco County Transportation Authority and PBConsult, San Francisco.

Cherchi, E., Polak, J. and Hyman, G. (2004) The impact of income, tastes and substitution effects on the assessment of user benefits using discrete choice models, presented to *European Transport Conference*, Strasbourg.

Cochrane, R. A. (1975) A possible economic basis for the gravity model, *J. Transport Economics and Policy*.

CPB and NEI (2000) *Evaluatie van infrastructuurprojecten: leidraad voor kosten-batenanalyse* (Evaluation of infrastructural projects: guide for cost-benefit analysis), CPB/NEI, The Hague/Rotterdam.

Daly, A. (2004) Properties of random utility models of consumer choice, presented to TraLog Conference, Molde.

Daly, A.J. and Zachary, S. (1976) Improved multiple choice models. In *Proceedings of the Fourth PTRC Summer Annual Meeting*. PTRC, London.

Deaton, A. and J. Muellbauer (1980) *Economics and Consumer Behavior*, Cambridge University Press, Cambridge.

Ecorys Transport and 4cast (2004) *Verbeteren van bepaling directe effecten uit LMS/NRM output, naar een efficiëntere en betere werking van modeloutput* (Improving the determination of direct effects from LMS/NRM output, towards a more efficient and better use of model output), eindrapport voor AVV, Ecorys, Rotterdam.

EXPEDITE Consortium (2002) *EXPEDITE Final Publishable Report*, Report for European Commission, DGTREN, RAND Europe, Leiden

Gupta, S., Kalmanje S., and Kockelman, K.M. (2004) *Road Pricing Simulations: Traffic, Land Use and Welfare Impacts for Austin, Texas*. Paper submitted for presentation at the 84th Annual Meeting of the Transportation Board and for Publication *Transportation Research Record*.

Harris, A. J. and Tanner, J. C. (1974) *Transport demand models based on personal characteristics*, TRRL Supplementary Report 65UC.

Herriges, J. A. and Kling, C. L. (1999) *Nonlinear income effects in random utility models*, *Review of Economics and Statistics*, 81 (1), pp. 62–72.

Johansson, P.-O. (1993) *Cost-benefit analysis of environmental change*, Cambridge University Press, Cambridge.

Jong, G.C. de, R. Plasmeijer, P. Sanders and P. Warffemius (2004) *De waarde van betrouwbaarheid van reistijden in personen- en goederenvervoer*, *Tijdschrift Vervoerwetenschap*, 40-4, pp. 29-34.

Kalmanje, S. and Kockelman, K.M. (2004), *Credit-Based Congestion Pricing: Travel, Land Value and Welfare Impacts*, Paper to be presented at the 2004 Annual Meeting of the Transportation Board and recommended for Publication in *Transportation Research Record*.

Karlström, A. (2000) *Non-linear value functions in random utility econometrics*, presented to IATBR Conference, Gold Coast, Australia.

Koopmans, C.C. and Kroes, E.P. (2004) *Werkelijke kosten van files tweemaal zo hoog* (The real cost of queues twice as high), article in *ESB April 2004*, pp. 154-155

Kouwenhoven, M., R. van Grol and E. Kroes (2004) *Ontwikkeling ex ante instrument betrouwbaarheid*; Report TR-233-AVV, RAND Europe, Leiden.

McFadden, D. (1978) *Modelling the choice of residential location*. In Karlqvist, A., Lundqvist, L., Snickars, F. and Weibull, J. (eds) *Spatial Interaction Theory and Residential Location*. North-Holland, Amsterdam.

McFadden, D. (1981) *Econometric models of probabilistic choice*. In Manski, C. and McFadden, D. (eds) *Structural Analysis of Discrete Data: With Econometric Applications*. The MIT Press, Cambridge, Massachusetts.

McFadden, D. (1996) *On computation of willingness-to-pay in travel demand models*, Dept. of Economics, University of California, Berkeley (unpublished).

Odeck, J., Rekdal, J. and Hamre, T.N. (2003) The Socio-economic benefits of moving from cordon toll to congestion pricing: The case of Oslo. Paper presented at *TRB Annual meeting*.

Raad, P.M. de (2004) Congestion costs closer examined, Monetarising adaptive behaviour using the Logsum-method, MSc Thesis, TU Delft.

RAND Europe (M. Pieters and B. Zondag) (2004) Voorstudie koppeling LMS en TIGRIS XL (Pre-study interaction between LMS and TIGRIS XL), commissioned by AVV, RAND Europe, Leiden.

RAND Europe (R. Hamer, G. de Jong en E. Kroes, P. Warffemius) (2005) The value of reliability in transport, Provisional values for The Netherlands, Report TR-240-AVV, RAND Europe, Leiden.

Small, K. A. and Rosen, H. S. (1981) Applied welfare economics with discrete choice models, *Econometrica*, 49, pp. 105-130.

Train, K. (2003) Discrete choice methods with simulation; Cambridge University Press, Cambridge

US Department of Transportation, Federal Highway Administration (2004?) Case Study: Sacramento, California Methodology Calculation of User Benefits, USDOT, http://www.fhwa.dot.gov/planning/toolbox/sacramento_methodology_user.htm.

von Haefen, R. H. (2003) Incorporating observed choice into the construction of welfare measures from random utility models, *J. of Environmental Economics and Management*, 45, pp. 145-165.

Williams, H.C.W.L. (1977) On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A*, 9 (3), pp. 285-344.